

## FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA  
CCCACGCGTCCGGGCCGGAGCAGCACGGCCGAGGACCTGGAGCTCCGGCTGCGTCTTCCCG  
CAGCGCTACCCGCCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTG  
CTGCTGCCGCCCCGCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT  
GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA  
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCGCTGCTGGAGATC  
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA  
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT  
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC  
CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA  
GGGCGACGGGTCTGCCGGTGCCACATGGGGTACCAGGGCCCCGCTGTGCACTGACTGCATGG  
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC  
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT  
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG  
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACAGTGCGAAGAGTGTGACTCCAGCTGTGTG  
GGCTGCACAGGGGAAGGCCCAGGAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA  
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAAACCTGTGTGAGGAAAAACG  
AAAAGTGTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCTTGACGGCTTCGAAGAAACG  
GAAGATGCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT  
GCCCTCCCGCGAAGACCTGTAATGTGCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC  
CGTGGAATAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC  
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA  
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG  
AGGAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC  
TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA  
TCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTC  
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATGGCCTGAAAT  
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG  
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGC  
ATCTCAATTAGTCAGCAACCCAGTTTT

## **FIGURE 2**

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEKTLISKYESSEIRL  
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLLKVCCSPGTYGPDCLACQGGSQRPCSG  
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESKTCSGLTNRDCGECEVGWVLDE  
GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT  
CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL

### **Signal peptide:**

amino acids 1-24

### **N-glycosylation sites.**

amino acids 190-194 and 251-255

### **Glycosaminoglycan attachment sites.**

amino acids 149-153 and 155-159

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 26-30

### **Casein kinase II phosphorylation sites.**

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343  
and 349-353

### **Tyrosine kinase phosphorylation site.**

amino acids 303-310

### **N-myristoylation sites.**

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and  
313-319

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 308-320

### **EGF-like domain cysteine pattern signature.**

amino acids 166-178

### **Leucine zipper pattern.**

amino acids 94-116

### FIGURE 3

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC  
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCAGCCGTCTAAACGGGAACA  
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG  
GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGCATGGCCCCGGAGGAGCGCCTTC  
CCTGCCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC  
CGGGCCGCGCAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA  
TAGGATTTGAAGAAGATATCCTGATTGTTTTAGAGGGGAAAATGGCACCTTTTACACATGAT  
TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTAC  
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCCTGTCTTGCGCTCCCTGGATA  
AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA  
TCAGTTGTTCAAGTTGGTTTTCCCATGTCTTGGAACACAGGATGGGGTGGCAGCATTGAAGT  
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT  
TCTTTAAACATGTCAACAAGCTGAGTGCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT  
GAAAGACGCATCTGCGAGTGCTCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG  
TACCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG  
GATTCTATGGAGTGAACGTGTGACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACC  
TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG  
CAAATGCCACAAACCCTGTGCAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT  
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT  
GGAACCTGCCATGAACCCAACAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA  
TAAAAGGTACGAAGCCAGCCTCATACTGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGC  
ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG  
**TGA**ACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCA  
TGTGTTGAATGTTCAAATAATGTTTATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT  
TCATTATAAATCACTGAGCTGATATTTACTCTTCCTTTTAAAGTTTCTAAGTACGTCTGTAG  
CATGATGGTATAGATTTTCTTGTTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA  
TCAGGTAAAATTTTTCAGTGTGTAGTTGGCAGATATTTTCAAATTAACAATGCATTTATGGT  
GTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAAT  
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT  
GTTTGTACATTTTAAAAAATTGCTCTTAATTTTAAACTCTCAATACAATATATTTTGACC  
TTACCATTATTCCAGAGATTCAGTATTAAAAAATAAATACTGTGGTAGTGGCATT  
AAACAATATAATATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTCAT  
TGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT  
ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAATAAATAAATAA  
AAAAAATAAATAAATAAATAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC  
CGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

## FIGURE 4

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094

><subunit 1 of 1, 379 aa, 0 stop

><MW: 41528, pI: 7.97, NX(S/T): 2

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEEPLYLWIDAHQARVLIGFEEDILIVSEGK  
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLL  
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGC  
RNGGFCNERRICECPDGFHGHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST  
TCFNNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC  
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERRDP  
PESNYIW

### **Signal peptide:**

amino acids 1-28

### **N-glycosylation site.**

amino acids 88-92, 245-249

### **Casein kinase II phosphorylation site.**

amino acids 319-323

### **Tyrosine kinase phosphorylation site.**

amino acids 370-378

### **N-myristoylation sites.**

amino acids 184-190, 185-191, 189-195, 315-321

### **ATP/GTP-binding site motif A (P-loop).**

amino acids 285-293

### **EGF-like domain cysteine pattern signature.**

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTTCGACAGCCAGGAGGCGGAGGCGCGCGGGGCCAGCCTGGG  
CCCCAGCCACACCTTCACCAGGGGCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC  
TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGG  
GAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA  
GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT  
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC  
TGCCTCGGCGTGCCACCCCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTTCGTATCTA  
TCCAGTCTTGGGAACGTACTGGGACAACCTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT  
GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG  
GAACCACAGCGCCTTCTGGGGCATGACCCTGGATTGAGGGCATTTCGCTACCGCCTGGGCACCA  
TCCGCCCATCTTCCTCGGTTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG  
GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA  
CCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT  
CAATCCATTCTCTGGGACACATGACGCCTGTCTGTGCGCCCCAGAACCTGCTGTCTTGTGAC  
ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG  
CCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC  
CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTCCGGGCAAGCGCCAGGCCACTGCC  
CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT  
CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA  
TGGAGGTGCATGAGGACTTCTTCCTATAAAGGAGGCATCTACAGCCACACGCCAGTGAGC  
CTTGGGAGGCCAGAGAGATAACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG  
AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC  
CAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGGCGTCAATGAGTGCGACATCGAG  
AGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTTCATCACTGAGGCTG  
CGGGCACCACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCAATG  
GGGCGGTGACCCAGCCTCGCCCGACAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAAT  
CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCTGGGAGCCGCGGGCAGGCGAGACTGGCG  
GAGCCCCCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG  
ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC  
CCCAATAACCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG  
TTGCCCAGGTTGGAGTGAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA  
AGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACACCTGGC  
TAATTTTTTGATTTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCCAGGCTGGTTTCGAACT  
CCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC  
ACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTTCTTTTCACTGTTTTAAAA  
TAAACCAAAGTATTGATAAAAAAAAAA

## **FIGURE 6**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDC  
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYPVLGTYWDNCNR  
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

**N-glycosylation site.**

amino acids 78-82, 161-165

**Casein kinase II phosphorylation site.**

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,  
411-415

**N-myristoylation site.**

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,  
269-275, 378-384, 442-448

**Amidation site.**

amino acids 26-30, 318-322

**Eukaryotic thiol (cysteine) proteases histidine active site.**

amino acids 398-409

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 7**

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTGTTGTCTCAAATCCA  
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC  
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG  
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC  
CCCCCTGTATGATGCACAGCCGAGCCATGGGTGCGGGCAAGCGCCAGGCCACTGCCCCTGCGC  
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC  
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG  
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG  
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

Sequence

## **FIGURE 8**

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT  
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG  
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG  
CAGCTTGAACTGCGTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT  
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT  
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG  
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG  
GCCCAGTGGGAGCCTGTCCTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT  
CTGACCCCTGTCCCCCACCCTGACCCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA  
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC  
CATGGCCCAGCATTTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT  
CCCTGCCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA  
GGGGACAGGCACTCAGGAGGGGCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA  
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA  
GGGGCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT  
AATAAACACCTGTTGGATAAGCCAAAAAA

## )

(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
-----	-----	-----	-----	-----	-----	-----	-----

**Signal peptide:**

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

## FIGURE 10

CCCACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCCCGCCTGCTGCCAACCTCACTCTGT  
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACC  
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA  
CGTGCAGGTCACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC  
TCATAGTGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAG  
TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG  
CGTGTTACCGGAGATCGTGCTGGAGAACTATAACGGCCTTCCAGAACGCCCGGCACGAGG  
GCTGGTTCATGGCCTTCACGCGGCAGGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAAC  
CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGC  
CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCACCCGCCGGACCAAGCGCACAC  
GGCGGCCCCAGCCCCCTCACGTAGTCTGGGAGGCAGGGGGCAGCAGCCCCCTGGGCGCCTCCC  
CACCCCTTTCCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC  
GAGGGAGGACCTGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTG  
CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT  
GAGTGTACCCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC  
CGCTGAAAGGTCAGCGACTGAAGGCCTTGACAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA  
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACCTCCTCCTGGCTAGACTGTA  
GGAAGGGACTTTTGTTTGTGTGTGTGTTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAATAG  
AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCAACCCCAACTCCCAGCCC  
CGGAATAAAACCATTTTCCTGC

## **FIGURE 11**

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENNYTAFQONARHEGWFMATRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

### **Signal peptide:**

amino acids 1-22

### **N-glycosylation site.**

amino acids 9-13, 126-130

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 60-64

### **Casein kinase II phosphorylation site.**

amino acids 65-69

### **Tyrosine kinase phosphorylation site.**

amino acids 39-48, 89-97

### **N-myristoylation site.**

amino acids 69-75, 188-194

### **Amidation site.**

amino acids 58-62

### **HBGF/FGF family signature.**

amino acids 103-128

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG  
GACAGCAGCAAAGAGGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC  
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT  
CATTTTGATTTTGTCTGTTTATTTTTTTTTTTCTTTTTCTTTTTCCCACCACATTGTATTTTAT  
TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCAGCCATGGGGCTTTTTTCTT  
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC  
CTAGTGTGTGCCGCTGCGACAGGAACCTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG  
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC  
TGGATTTCTCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC  
AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC  
AATATTCAGACCATTTACGGGCTGCTCTTGCCAGCTCTTGAAGCTTGAAGAGCTGCACCT  
GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC  
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCTGTGGAC  
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTTCATATCCGACATGGCCTTCCAGAA  
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG  
AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC  
CACCTCCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT  
AAACCACATTCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA  
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC  
ACTGCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA  
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG  
GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCTGTCCCACCACGACCCCCGGCCTG  
CCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT  
TCCAAACCCTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAACTTCCCACGATTC  
CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC  
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTCACCGTGATGGCATA  
CAAACCTCACATGGGTGAAAATGGGGCCACAGTTTAGTAGGGGGCATCGTTTCCAGGAGCGCATAG  
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT  
TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC  
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA  
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT  
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTGCTGGCATATGCACAAAAGGGGCGCTACACCTC  
CCAGAAGTGGAATAACAACCGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA  
AGGACAACCTCCATCCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAAATAACGAT  
CAACTCCTTAAAGGAGATTTTCAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA  
CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC  
ACTGCCATACGTAAGCAGCCAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA  
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT  
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG  
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTTGTAACCTTTTGCTTTTTTAAATCTT

## **FIGURE 13**

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP  
EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI  
SRAALAQLLKLEELHLDNDSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR  
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD  
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN  
PWFCDCSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMVRELNMNLLSCPTTTPGLPLFTP  
APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND  
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL  
DAFNRYAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL  
LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG  
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

**Signal peptide:**

amino acids 1-42

**Transmembrane domain:**

amino acids 542-561

**N-glycosylation site.**

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

**Casein kinase II phosphorylation site.**

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

**Tyrosine kinase phosphorylation site.**

amino acids 319-328

**N-myristoylation site.**

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,  
522-528, 545-551, 633-639

**Amidation site.**

amino acids 581-585

**Leucine zipper pattern.**

amino acids 164-186

**Phospholipase A2 aspartic acid active site.**

amino acids 39-50

[illegible]

ACTTTGGAGCAACGCGCGCGCGGCGGAGACAGAGGCGAGAGGCGAGAAGCTGGGGCTCCGTCTCGCCTCCCCACGAGCG  
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCGTGCC  
TCGCTTCCCAGGCGCCGGCGGCTGCAGCCTTGGCCCTCTTGCTCGCCTTGAAAATGGAAAAGATGCTCGCAGGCT  
GCTTTCTGCTGATCCTCGGACAGATCGTCTCTCCCTGCCGAGGCCAGGGAGCGGTACGTTGGGAGGTCCATCT  
CTAGGGCGAGACACGCTCGGACCCACCCGAGACGGCCCTTCTGGAGAGTTCTCTGTGAGAACAAAGCGGGCAGACC  
TGTTTTTCATCATTGACAGCTCTCGAGTGTCAACACCCATGACTATGCAAAAGGTCAAGGAGTTTCATCTGTGGACA  
TCTTGCAATTCTTGGACATTGGTCTGTATGTACCCGAGTGGGCTGTCTCAATATGGCAGCATGTCAAGAATG  
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTGTCAAGAGGATGCGGCATCTGTCCACG  
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCCGGCCCTGA  
GGGAGAATGTGCCACGGGTATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA  
AGGCACGGGACACGGGCATCTAATCTTTGCCATTGGTGTGGGCGAGGTAGACTTCAACACCTTGAAGTCCATTG  
GGAGTGAGCCCCATGAGGACCATGTCTTCTTGTGGCCAATTTACGCCAGATTGAGACGCTGACCTCCGTGTTCC  
AGAAGAAGTTGTGCACGGCCACATGTGCAGCACCTTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG  
GCTCATACGTCTGCAGGTGCAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGAGAATCCAGGATCTGT  
GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG  
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAAAACCCAGGATGTGAAC  
ATGATGTGTAATGTCTAGTGGCTCTACCTTTGGCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAA  
CGTGCACAAGGATCAACTACTGTGCATGAACAAACCGGGCTGTGAGCATGAGTGCCTCAACATGGAGGAGCT  
ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC  
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCTTTCGTCTGCCAGTGCTCAGAAGGCTTCC  
TCATCAACGAGGACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGCTGAGTGACCATGGTTGTGAATACTCCT  
GTGTCAACATGGACAGATCCTTTGCTGTCACTGTCTGAGGGACAGTGCTCCGCAGCGATGGGAAGACGTGTG  
CAAAATTTGACTCTTGTGCTCTGGGGGACCAACGGTTGTGAACATTCGTGTGAAGCAGTGAAGATTCGTTGTGT  
GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAGATGTCTGCCAAGCTATAG  
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACAGTGCGAGTGCTTGGAGGGATTCCGGCTCG  
CTGAGGATGGGAACCGCTGCCGAAGGAAGGATGTCTGCAATCAACCCACCATGGCTGCGAACACATTTGTGTTA  
ATAATGGGAATTCCTACATCTGCAAAATGCTCAGAGGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT  
GCACTGAAGGCCCAATTGACCTGGTCTTTGTGTAGTATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTGAGGTG  
TGAAGCAGTTTGTCACTGGAATTATAGATTCTTGAACAATTTCCCCAAAGCCGCTCGAGTGGGCTTGTCCAGT  
ATTCACACAGGTCCACACAGAGTTCACTCTGAGAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCC  
ACATGAAATACATGGGAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTACCCAAG  
GAGAAGGGGCCAGGCCCCCTTCCACAAGGGTGCCACGAGCAGCCATTGTGTTCCACCGACGGACGGGCTCAGGATG  
ACGCTCTCCGAGTGGGCGAGTAAGCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG  
AGGAGGAACACAAGAGATTGCTCTGAGCCCAAAACAAGCATCTCTTCATGCCGAAGACTCAGCACAATGG  
ATGAGATAAGTGA AAAA ACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG  
CAGGGGAAC TGCCAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCAACATAAATATCCAAGACCTACTTT  
CTGTTCTAATTTTGAGTGTCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT  
CCCATCTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACCGATCAATGCAATGTGAAAACCTTATAATGT  
TCCAGAACCTTGCAACGAAGAAGTAAAGAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC  
TGGA AAATCGCCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTATCAGGATTACAAT  
GAACGCAGTGCAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAACAATCAGTACTGA  
GAAACCTGGTTTGCCACAGAACAAAGACAAGAAGTATACTAACTTGTTATAAATTTATCTAGGAAAAAATCCT  
TCAGAATCTAAGATGAATTTACAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAAATATACTGTGGACAC  
AATTGCTTCTCGCTCATCTTGCTTGTAGTGCAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTT  
CTGTAGAACACTGGCCATAGGAATGCTGTTTTTTGTACTGGCACTTACCTTGATATATGTTATGGATGTATG  
CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAATAATACCACCTTCAG

## **FIGURE 15**

MEKMLAGCFLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFII  
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV  
KMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD  
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN  
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA  
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKP GC  
EHECVNMEESSYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGF LI  
NEDLKTC SRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTC AKLDS CALGDHGCE  
HSCVSSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA  
EDGKRCRRKDVCCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID  
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA  
HMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN  
GITMYAVGVGKAIEEELQEIASEPTNKHLYAEDFSTMDIEISEKLKKGICEALEDSDGRQDS  
PAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL  
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQORMEALENRLRYR

**Signal peptide:**

amino acids 1-23

**N-glycosylation site.**

amino acids 221-225

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 115-119, 606-610, 892-896

**Casein kinase II phosphorylation site.**

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,  
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,  
781-785, 819-823, 866-870

**N-myristoylation site.**

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,  
639-645, 690-696, 752-758, 792-798

**Amidation site.**

amino acids 314-318, 560-564, 601-605

**Aspartic acid and asparagine hydroxylation site.**

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,  
540-546, 581-587

## FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTGCGCGAGCCTCG  
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGC  
CATGATTTCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA  
GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCCGCAACCGGTTGCAG  
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGGGGAGGTGTCTTC  
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC  
AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC  
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA  
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT  
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT  
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA  
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG  
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC  
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC  
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC  
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC  
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT  
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT  
TGACCCCCACGCCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT  
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTTCTTCCTCTGGCTTGAGCCG  
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCAGAGTCAAGCTGGCTCTCTGGTATTGATGAC  
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCCTATAAGGGTCACCTCTAGCAC  
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC  
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA  
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG  
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC  
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT  
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC  
TTATTAATAAATAACATGAAATATGTGTTGTTTTCATTTGCAAATTTAAATAAAGATACATAA  
TGTTTGTATGAAAAA

## **FIGURE 17**

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS  
SQPWEVPPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY  
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSQCSPRSKPAVQYQ  
WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA  
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTL  
SSVTSARALRPPHGP RP GALTP T PSLSSQALPSPRLPTTDGAHPQPI SPIPGGVSSSGLSR  
MGAVPVMVPAQSQAGSLV

### **Signal peptide:**

amino acids 1-29

### **Transmembrane domain:**

amino acids 245-267

### **N-glycosylation site.**

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

### **N-myristoylation site.**

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,  
262-268, 308-314, 363-369, 364-370

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 164-175

[illegible]

CGCCACCACTGCGCCACCACCAATGAAACGCCTCCCGCTCTAGTGGTTTTTCCACTTTG  
TTGAATTGTTCTCTATACCTCAAATTTGCACCAAGACACCTTGCTCTCCCAAATGCAAAATGTGA  
AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCAAA  
TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCTGTGGCGAAATGCTAATTGC  
ACTAACACAGAAGGAAGTTATTATTGTATGTGTACCTGGCTTCAGATCCAGCAGTAACCA  
AGACAGTTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACCTGCCATT  
TAGATAATGTCGTATAGCTGCAAATATTAATAAAACCTTAACAAAAATCAGATCCATAAAA  
GAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATCTGTGACAGATCTTTCACCAACAGA  
TATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAA  
CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT  
AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC  
ACATCTTACAAAACCTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCAGAGCTTCC  
AAAAGACCACAGAGTTTGATACAAATTC AACGGATATAGCTCTCAAAGTTTTCTTTTTTGAT  
TCATATAACATGAAACATATTATCCTCATATGAATATGGATGGAGACTACATAAATATATT  
TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATA  
AGAGTATTGGTCCTTTGCTTTTCATCATCTGACAACCTCTTATTGAAACCTCAAATTATGAT  
AATTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCC  
ACCCACATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA  
GGTATAGGAGTCTATGTGCATTTTGGAAATTAACCTGATACCATGAATGGCAGCTGGTCT  
TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT  
GACACATTTTTGCAATTTTTGATGTCCTCTGGTCTTCATTGGTATTAAGAGATTATAATATTC  
TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT  
ACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG  
TAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCT  
TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTGCATGGATGTGC  
ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTTTGCA  
CAAGAATTTTTATATCTTTGGCTATCTAAGCCAGCCGTGGTAGTTGGATTTTCGGCAGCAC  
TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACCTTTATT  
TGGAGTTTTATAGGACCAGCATGCCTAATCATCTCTGTTAATCTCTTGCTTTTGGAGTCAT  
CATATACAAAGTTTTTTCGTCACACTGCAGGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACA  
TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCTTCTCGGCACCACCTGGATCTTT  
GGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC  
TTTCCAGGGGATGTTCATTTTTTTATTCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAAT  
ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAGGTAAACATAGAGAATG  
GTGGATAATTACAACCTGCACAAAAATAAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA  
TGACTCATCAAATTATCCAATTATTAACCTACTAGACAAAAAGTATTTTAAATCAGTTTTTCT  
GTTTATGCTATAGGAACTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGT  
TTTTCTATGTGAAATAGTTCTGTCAAAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTT  
CTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTTCTAACACGAGAAGTATATGAA  
TGTCTTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTGCTTTGAAACTAGTCC  
CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGA  
ATATCAAACAGTGAAAAGGGAATGATAAGATGTATTTTGAATGAACTGTTTTTTCTGTAGAC  
TAGCTGAGAAATTGTTGACATAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA  
TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGTTTCTT  
TTTCTAATATTCTAAAAAATAAAGGTTTACCTCCACAAATTGAAAAAATAAAGGTTTCTT  
AAAAAATAAAGGTTTACCTCCACAAATTGAAAAAATAAAGGTTTCTT

## **FIGURE 19**

MKRLPLL VVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC  
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVC IENVNANCHLDNVCIAA  
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL  
SNSTLT E FVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT  
NSTDIALKVFFFDSYNMKHIHPHMNDGDYINIFPKRKAAYDSNGNVAVAFLLYYKSIGPLLS  
SSDNFLLKPQNYDNSEEEERVISVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF  
WNYSPTMNGSWSSEGC E LTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG  
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFC SIIAGL  
LHYFFLA AFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFS AALGYRYYGT  
TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA  
LALLFLLGTTWIFGVLVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV  
PCCFGCLR

### **Signal peptide:**

amino acids 1-19

### **Transmembrane domain:**

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,  
648-664

### **N-glycosylation site.**

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,  
188-192, 249-253, 381-385, 395-399

### **Glycosaminoglycan attachment site.**

amino acids 49-53

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 360-364

### **Casein kinase II phosphorylation site.**

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,  
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,  
346-350, 608-612

### **Tyrosine kinase phosphorylation site.**

amino acids 36-44, 669-677, 670-678

### **N-myristoylation site.**

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,  
434-440, 480-486, 521-527

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 75-87

## **FIGURE 20**

TGGAAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAATATATTTCCAAAGNG  
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATAAGAGTAT  
TGGTCCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAATTATGATAATTCT  
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTAGTCTCAATGAGCTCAAACCCACCCAC  
ATTATATGAACTTGAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATAGGTATA  
GGAGTCTATGTGGCATTTTGGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG  
GGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA  
TTTTGCAATTTTGATGTCCTCTGGTCCTTCATTGGTATTAAAGATTATAATATTCTTACAA  
GGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTTTACCTTC  
TGGTTCTTCAGTGAAATTCAAAGCACCAGGA

U.S. Patent & Trademark Office

## FIGURE 21

GCTCCCAGCCAAGAACCTCGGGGCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG  
CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCCGGCTGGGA  
CAAGAAGCCGCGCCTGCCTGCCCGGGCCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG  
GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG  
TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG  
CCGCGCGCCGTGAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC  
GGCACTCACAGCCCCGAGCGCATCCCGGTGCGCGCCAGCCTCCCGCACCCCCATCGCCGG  
AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG  
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC  
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC  
ACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCCGACGGCGTCTGGACTGCGCGCGGGGC  
CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG  
CGTGACAGCGTGCGGTACCTCTGCATGGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT  
ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA  
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA  
CAGAGGCTTTCTTCCACTCTCTCATTTTCTGCCATGCTGCCCATGGTCCCAGAGGAGCCTG  
AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCTGGAGACCGACAGCATG  
GACCCATTTGGGCTTGTACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAAGTAACT  
GAGACCATGCCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG  
TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA  
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT  
CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTTCTGCTCCCTCGA  
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC  
TCACTTCTTTTGGAATAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC  
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTAGGAACAGGTGATCCACTCTGTA  
AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG  
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCCACCAG  
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC  
CTGAGGCCAGTTCTGTCATGGATGCTGTCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC  
TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT  
CCCAGGCCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT  
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC  
CCTTTCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA  
TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATATTTATGTATGTAAGTGAGGTTTG  
TTTTGTATATTAAAATGGAGTTTGTGTTGT

## **FIGURE 22**

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLRI  
RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVH SVRYLCMGADGKMQGLLQYSEEDCAFEE  
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD  
MFSSPLETDSMDPFGLVTGLEAVRSPSF EK

### **Signal peptide:**

amino acids 1-22

### **Casein kinase II phosphorylation site.**

amino acids 78-82, 116-120, 190-194, 204-208

### **N-myristoylation site.**

amino acids 15-21, 54-60, 66-72, 201-207

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 48-59

FIGURE 22

## FIGURE 23

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA  
GAGCAGCCGGCTGCCGCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG  
CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA  
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA  
AGAAGACTGTTTTCTCCAGATTAGAGTGGAAGAACTGGGTTCGGAGTGTCTCCTTTGTCTAC  
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG  
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG  
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT  
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAAGTGTGGTAGAGCTACGATGTCAAGA  
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA  
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATAACAATGAATACAAAACCTGGAAGT  
CTGCAATTTAATACTGTTTCCAACTGGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC  
TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA  
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT  
GCTCAGAGGAAAGGCTACTTTTCAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA  
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG  
CCGCGGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC  
CATCTCTACTAAAATACAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCAGCTGC  
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC  
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCCATCTCAAAAATAAAATAAATAAATA  
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TABLE 1. *Continued*

**Signal peptide:**

**Transmembrane domain:**

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Amidation site.

amino acids 226-230

## FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG  
 AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT  
 CATGCTGCTATTCTCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA  
 ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA  
 TTATATCATTAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTGGACAATG  
 CAATTGTGGCACTGGCACTTATTTTCAGTGAAGAAAACTTTGTGGTTCTATGGCATTCATCA  
 TTTGACAAATGCAAGCATCTTCCTTATCAATCAGCTCCTATTGAACCTTACTAGCACTGACTG  
 TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT  
 CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG  
 TGGATTGTCCACGGTTATGTACGTGTGAATCAGGCCTTGGTTTACACCCAGATCCATTTAT  
 ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCAGCCAGATTGCC  
 AGCTAACACACAGATTCTTCTCCTACAGACTAACAAATATTGCAAAAATTGAATACTCCACAG  
 ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT  
 ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAACTTACTGA  
 ACTGCCTGAAAAATGTCTGTCCGAAGTGAAGCACTTACAAGAACTCTATATTAATCACAACT  
 TGCTTTCTACAATTTACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC  
 AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT  
 TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACCTTTAAGCCTCTTATCA  
 ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT  
 GGACTGGAAAACCTTAGAAAGCATCTCTTTTACGATAACAGGCTTATTAAAGTACCCCATGT  
 TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTGGATCTAAATAAAAAATCCTATTAATAGAA  
 TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAAGAGTTGGGGATAAATAATATGCCT  
 GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC  
 TACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG  
 AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG  
 CCAAACCTCAAGGAAATCAGCATAACAGTAACCCCATCAGGTGTGACTGTGTGCATCCGTTG  
 GATGAACATGAACAAAACCAACATTCGATTCTATGGAGCCAGATTCAGTGTGTTTGGCGTGACC  
 CACCTGAATTCGAAGGTGAGAATGTTTCGGCAAGTGCATTTTCAGGGACATGATGGAAATTTGT  
 CTCCCTCTTATAGCTCCTGAGAGCTTTTCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT  
 TTCCTTTCACTGTAGAGCTACTGCAGAACACAGCCTGAAATCTACTGGATAACACCTTCTG  
 GTCAAAAACTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA  
 GATATAAATGGCGTAACTCCCAAAGAAGGGGGTTTATATACTTGTATAGCAACTAACCTAGT  
 TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG  
 GCTCTTTGAATATTAAAAATAAGAGATATTGAGGCCAATTCAGTTTTGGTGTCTTGGAAAGCA  
 AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA  
 TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC  
 CATCAACTGAGTATAAAATTTGTATTGATATTCCCACCATCTATCAGAAAAACAGAAAAAAA  
 TGTGTAAATGTCACCACCAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC  
 CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGTCTTATCA  
 GCTGCCTCTCTCCAGAAATGAACTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG  
 AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA  
 AGAAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCTT  
AAAAACCACCAAGGAAACCTACTCCAAAAATGAAC

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

## **FIGURE 26**

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT  
FPARLPANTQIILLQTNNAKIEYSTDFVNLTLGLDLSQNNLSSVTNINVKKMPQLLSVYLE  
ENKLTPEKCLSELSNLQELYINHNLSTISPGAFI GLHNLRLHLNSNRLQMINSKWFDA  
LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL  
IKVPHVALQKVNLKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLDP  
LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSPNIRC  
DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGNVRQVHFRDMMEICLPLIAPESFPSNLNV  
EAGSYVSFHCRTAEAPQPEIYWITPSGQKLLPNTLTDFYVHSEGTLDINGVTPKEGGLYTC  
IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFAV  
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE  
YEKNNTTTLMACLGGLLGIIGVICLISCLSPENMCDGGHSYVRNYLQKPTFALGELYPPLIN  
LWEAGKEKSTSLKVKATVIGLPTNMS

### **Signal sequence:**

amino acids 1-22

### **Transmembrane domain:**

amino acids 633-650

### **N-glycosylation site.**

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,  
608-612, 624-628, 625-629

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

### **Tyrosine kinase phosphorylation site.**

amino acids 570-579

### **N-myristoylation site.**

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,  
491-497, 492-498, 634-640, 702-708

### **Cell attachment sequence.**

amino acids 277-280

## **FIGURE 27**

GCCCCGGGACTGGCGCAAGGTGCCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTAG  
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC  
CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT  
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA  
TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG  
GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA  
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTAAAG  
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT  
GCCTTCAAAGGAGTAGCTGAAACCTTGACAGACTCTGGACTTGTCGACAATCGGATTCAAAG  
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC  
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC  
AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCTCAATGCTGC  
CAACGACGCTGACCTTTGTAACTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCA  
TGTTTGGCTGGTTCACCTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG  
GATGCCCCGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA  
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAACTGACTGTCATTGAGAAAGAAAGAAA  
GTAGTTTGCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA  
ACTTTGTATTTTCAAGTTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA  
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGT  
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT  
AATTTAAAAGCAAATAAAAGCTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAACA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 28**

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL  
PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDSLSDNR  
IQSVHKNAFNNLKHARARIANNPWHCCTLQQVLRSMASNHETAHNVICHTSVLDEHAGRPFL  
NAANDADLCNLPKKTDDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK  
ADEPDDISTVV

### **Signal sequence:**

amino acids 1-33

### **Transmembrane domain:**

amino acids 205-220

### **N-glycosylation site.**

amino acids 47-51, 94-98

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 199-203

### **Casein kinase II phosphorylation site.**

amino acids 162-166, 175-179

### **N-myristoylation site.**

amino acids 37-43, 45-51, 110-116

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATG CAGGTGAGCAAGAGGATGCTGGCGGGG  
GGCGTGAGGAGCATGCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT  
GGGCTCAGTGCTGTCAGGCTCGGCCACGGGCTGCCCCCGCTGCGAGTGCTCCGCCCAGG  
ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG  
ACGCGCCTGCTGGACCTAGGCAAGAACC GCATCAAAACGCTCAACCAGGACGAGTTCGCCAG  
CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCCGGCG  
CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAAGCTCATC  
CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT  
CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG  
ACAATGACCTCGTCTACATCTCTCACC GCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG  
ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCACCTGCACGGCCT  
CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC  
TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC  
TGCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC  
CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCCTCAACCTCTCCTACAACCCCATCA  
GCACATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC  
GGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCGCGGGCCTCAACTACCTGCGCGTGCTCAA  
TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG  
AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGG  
CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCACGTGCGCCACGCCCCGAGTTTGTCCA  
GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCGCGCGCG  
CCCGCATCCGGGACCGCAAGGCCAGCAGGTGTTTGTGGACGAGGGCCACACGGGTGCAGTTT  
GTGTGCCGGGCCGATGGCGACCCGCCGCCCATCCTCTGGCTCTCACCCCGAAAGCACCT  
GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT  
ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGAC  
TCCATGCCCCGCCACCTGCATGTGCGCAGCTACTCGCCCCGACTGGCCCCATCAGCCCAACAA  
GACCTTCGCTTTTATCTCCAACCAGCCGGGCGAGGGAGAGGCCAACAGCACCCGCGCCACTG  
TGCCTTTCCCTTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC  
CTGGGCGTCGTCCTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC  
AAAGCACAAATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG  
ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCCGGGGCGGGGGGCAGGGACCCCCG  
GGCGGCCGGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCTTCCCACCTC  
CTCCCTACCCTTCTACACACGTTCTCTTTCTCCCTCCCGCCTCCGTCCCCTGCTGCCCCCG  
CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA  
CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA  
ATAATTCAATAAAAAAGTTACGAAC TTTCTCTGTAACCTTGGGTTTCAATAATTATGGATTTT  
TATGAAAACCTGAAATAATAAAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 30**

MQVSKRMLAGGVRSMPSPLLACWQPIILLVLGSLVSGSATGCPPRCECSAQDRAVLCHRKCF  
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFNNLFLNLRTL  
GLRSNRLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFDLYNLKSLEVGDNDLVYISHRA  
FSGLSLEQLTLEKCNLTSLIPTALSHLHGLIVLRLRHLNINAI RDYSFKRLYRLKVLEISH  
WPYLDTMPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRLNLSYNPISTIEGSM LHEL  
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSGNLETLILDSNPLA  
CDCRLLWVFRRRWRLNFNRRQOPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV  
FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGR LTVFPDGTLEVRYAQVQDNGTYL  
CIAANAGGND SMPAHLHVRSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI  
IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

### **Signal sequence:**

amino acids 1-41

### **Transmembrane domain:**

amino acids 556-578

### **N-glycosylation site.**

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,  
492-496, 505-509, 526-530, 542-546

### **Casein kinase II phosphorylation site.**

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,  
607-611

### **Tyrosine kinase phosphorylation site.**

amino acids 590-598

### **N-myristoylation site.**

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,  
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 32-43

Downloaded from www.sciencedirect.com

## FIGURE 31

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAAC  
ATCGTAGTCCACCCCCTCCCCATCCCCAGCCCCGGGGATTAGGCTCGCCAGCGCCCAGCC  
AGGGAGCCGGCCGGGAAGCGCGATGGGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGC  
TGTTTCGCTGCTGCTGGGCGCCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCTGG  
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCCTGGTGGTCAAGTGCCAAGTGAAAGATCA  
CGAGGACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA  
GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC  
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT  
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCACAGAAGCCCATCATCACTGGTT  
ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAG  
CCTGCAGCCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCAT  
ACAGGAAGATCCCAATGGTAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC  
GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC  
AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC  
TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC  
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT  
GCCCTGATCTTCCCTTTCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG  
CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCT  
CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTG  
CTCATCATGCTCATCTTCCTTGGCCACTACTTGATCCGGCACAAGGAACCTACCTGACACA  
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG  
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTTCATCTAGAGGCGCCTGCCCCACTTCCTGC  
GCCCCCAGGGGGCCCTGTGGGGACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAA  
CCGCAGGGCCGCCCCCTCCCGCTTGCTCCCCAGCCCACCCACCCCTGTACAGAATGTCTGC  
TTTGGGTGCGGTTTTGTACTCGGTTTGGAATGGGGAGGGAGGAGGGCGGGGGAGGGGAGGG  
TTGCCCTCAGCCCTTTCCTGCTGCTCTGCAATTTGGGTTATTATTATTTTTGTAAACAATCC  
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACA  
AACAAAAACA

## **FIGURE 32**

MGAPAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDET VVAGGTVVLKCQVKDHEDSSLQW  
SNPAQQOTLYFGEKRALRDNRIQLVTSTPHELSSISISNVALADEGEYTC SIFTMPVRTAKSLV  
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRI QEDPNGK  
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQR IEVLYTPTAMIRPDPPHPREGQ  
KLLHCEGRGNPVPVQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATS NMGSYKA  
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLIMLI FLGHYLIRHKGTYLTHEAKGSDD  
APDADTAIINAEGGQSGGDDKKEYFI

### **Signal sequence:**

amino acids 1-20

### **Transmembrane domain:**

amino acids 331-352

### **N-glycosylation site.**

amino acids 25-29, 290-294

### **Casein kinase II phosphorylation site.**

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

### **N-myristoylation site.**

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,  
306-310, 334-340, 360-364, 385-389, 386-390

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

FIGURE 32

## FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCTTTTCTTCTCCTTTCTG  
CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG  
TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT  
GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA  
TCTGTTCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAGGGCTTCACA  
AGTCTGCAGCGTTTTCACTGCCCCGACTTCCCAGTTTTTACCATTTATTTCTGCATGGCAATTC  
CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG  
AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG  
CTGCACATCAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA  
TCTGGAATATCTCCAGGCTGATTTTAAATTTATTACGAGATATAGACCCGGGGGCTTCCAGG  
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCTACCTGCCAAC  
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC  
CTATGAGGAGGTCTTGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCTT  
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC  
CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC  
CACCGAACAGGACTTGTGTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG  
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG  
GATCATGCCACACCAGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT  
CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCTTAGCTAACA  
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC  
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA  
GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA  
ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG  
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA  
GAAATTTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA  
TCCTCCCGGGCACTTTCAATGCCATGCCCAAACCTGAGGATCCTCATTCTCAACAACAACCTG  
CTGAGGTCCCTGCCTGTGGACGTGTTCGCTGGGGTCTCGCTCTCTAACTCAGCCTGCACAA  
CAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA  
TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA  
GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACCTTCTT  
TAGAAAGGATTTTCATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT  
CGCCACGTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC  
AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT  
GTTTGTACCTCCGCCTTACCCTGGTGGGCATGCTCGTGTATTATCCTGAGGAACCGAAAGC  
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT  
GACTCTTCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCACAGAGTGATGACTG  
TGGCTCTCACTCGCTCTCAGACTAAGACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG  
ATACATCCTTCCCCACCGCAGGCACCCCGGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG  
CCATCAGCTGGATGGGCATAAGTAGATAAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT  
GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA  
GCCAGCTCGCTCTTTGCTGAGAGCCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG  
AACTGACAGTGCCCTCGCCCTCGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC  
ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCTGTGGATTAG  
CCCCGTGATGGCTCCTGTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT  
TGTAATAAGTAACTTTGACTTCTGAC

## **FIGURE 34**

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH  
LFLHGNSLTRLPNEFANFYNAVSLHMENGLHEIVPGAFLGLQLVKRLHINNKNIKSFRKQ  
TFLGLDDLEYLQADFNLLRDIDPGAQDLNKLEVLILNDNLISTLPANVFQYVPITHDLRG  
NRLKTLPEYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ  
GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTK  
IPGNWQIKIRPTAAIATGSSRNKPLANSPLCPGGCSCDHIPGSGLKMNCCNNRVSSLADLKP  
KLSNVQELFLRDNKHISIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLDDLRLWLYMDSNY  
LDTLSREKFAGLQNLLEYLNVYNAIQILIPGTFNAMPKLRLILNLLRSLPVDVFAGVSL  
SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSTIVPFKQWAERLGSEVLMSDLKC  
ETPVNFFRKDFMLLSNDEICPOLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL  
VPGLLLVFVTSFTVVGMLVFILNRNRKRSKRDRANDSSASEINSLQTVCDSSYWHNGPYNADG  
AHRVYDCGSHSLSD

### **Signal sequence:**

amino acids 1-15

### **Transmembrane domain:**

amino acids 618-638

### **N-glycosylation site.**

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 122-126, 646-650

### **Casein kinase II phosphorylation site.**

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,  
608-612, 657-661, 666-670, 693-697

### **N-myristoylation site.**

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,  
354-360, 465-471, 493-499, 598-604, 603-609

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 337-348

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 35

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCCAGAATAACTCAGGGC  
TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCTGTGCGCGCCTAAGGGAAACTGTTGGC  
CGCTGGGCCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTGCGGAGCGAGGGCGGAGGGG  
AAGGGAGGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC  
AGCTCTGCGTCTCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCCTCAG  
AGAATGAGGCCGGCGTTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCCGGGCGGGCGG  
CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGC  
ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC  
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG  
GCCCCGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTGCGACTGGAGCGCAGGCGTTCCCACT  
GCACCCTGGAGAACGAGCCTTTGCGGGGTTTTCTCCTGGCTGTCTCCGACCCCGGCGGTCTC  
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC  
GGTACTCCAGGCCACCGGTGGGGTTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC  
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCTGCGCCGCGCCCCGGG  
GCCGCTCTAACTTGAGCTATCGCGCGCCCTTCAGCTGCACAGCGCCGCTCTGGACTTCAG  
TCCACCTGGGACCGAGGTGAGTGCCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA  
TCGCGGACGAAATCGGCGCTCGCTGGGACAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC  
GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG  
CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA  
GTGGGGAAGGACAGCCGACCCTTGGGGGGACCGGGGTGCCACCAGGCGCCCGCGGCCACT  
GCAACCAGCCCCGTGCCGAGAGAAATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC  
ACCACTTGTCCTGAACAAGACAATTCAGTAACATCTATTCTGAGATTCTCGATGGGGAT  
CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC  
CCATCAGGGAGCGTGATTTCCAAGTTTAAATTTCTACGACTTCTCTGCCACTCCTCAGGCTTT  
CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTGGTGATCT  
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTACGAAAGCCCCCTCTTCCCAGCCA  
AGGAAGGAGTCTATGGGCCCCGCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC  
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG  
CAGAGGGTGCTTGTGTCGGAGTCCCCCTCTTGGCTCTAGTGATGCATAGGGAAACAGGGGA  
CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTAC  
TTGTGTAAGTACAATTTCTGCAGAAATCCCCCTTCTCTAAATTCCCTTTACTCCACTGAG  
GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA  
TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA  
TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAACAAATAGAACACAATATAATTTACA  
TTAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTTCAAGGCTAGGAGTAT  
ATTGGTTCGAAATCCCAGGGAAAAAATAAAAAATAAAAAATTAAAGGATTGTTGAT

## **FIGURE 36**

MRPAFALCLLWQALWPGPGGGGEHPTADRAGCSASGACYSLHHATMKRQAEEACILRGGALS  
TVRAGAE LRAVLALLRAGPGPGGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE  
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGA  
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG  
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPPATA  
TSPVPQRTWPPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP  
SGSVISKFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILTM TVLGLVKLCFHESPSSQPR  
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 399-418

### **N-glycosylation site.**

amino acids 189-193, 381-385

### **Glycosaminoglycan attachment site.**

amino acids 289-293

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 98-102, 434-438

### **Casein kinase II phosphorylation site.**

amino acids 275-279, 288-292, 342-346, 445-449

### **N-myristoylation site.**

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,  
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,  
477-483

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 262-274

## FIGURE 37

CGGACGCGTGGGATTTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG  
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGGCCGCGTATC  
CCCCGGCTACCTGGGCCGCCCCGCGCGGTGCGCGCGTGAGAGGGAGCGCGCGGGCAGCCGA  
GCGCCGGTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTGGGTGCGGA  
GGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAACCCCGAGCGTCTACGCTGCCATGA  
GGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG  
CAGCAGTCCCCAGAGAGACCTGTTTTTACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT  
TATTGGCAGTGAAGGTTTTCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAAAATCA  
CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC  
CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTTGGCCG  
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA  
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA  
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAC  
CCCCAACTGGCCAGACCGGGATTACCCTGCAGGAGTCACTTGTGTGTGGCACATTGTAGCCC  
CAAAGAATCAGCTTATAGAATTAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC  
CGATATGATTATGTGGCTGTGTTTAAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA  
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT  
TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA  
AAACTGCCTACAACACAGAACAGCCTGTCAACCACCACATTCCCTGTAACCACGGGTTTAAA  
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT  
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAAACCATCACTCGCGATGGGAGTTTG  
CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTTCAGCAGGCGGG  
CAAGAACATGAGTGCCAGGCTGACTGTGCTCTGCAAGCAGTGCCTCTCCTCAGAAGAGGTC  
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC  
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG  
TTAAACAGTGAAGTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTT  
TCTCAGTAGAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAGATGG  
GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT  
TGCCTGCTGTGAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA  
AGCTAAAAGTGTCAAGCGTTGACAGCTTGGAAAGCGTTTATTTATACATCTCTGTAAAAGGAT  
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAGATTTTAGAAGTGCAATATTTATAGT  
GTTATTTGTTTACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA  
AATCAATGCTTAATAAAATATTTTAAAGGAAAAAAAAAAAAA

## **FIGURE 38**

MRGANAWAPLCLLLAAATQLSRQQSPERPVFCTCGGILTGESGFIGSEGFPGVYPPNSKCTWK  
ITVPEGKVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV  
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV  
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPAPIVSERNELLI  
QFLSDLSLTADGFIGHYIFRPKKLPPTTTEQPVTTFPVTGLKPTVALCQQKCRRTGTLEGN  
YCSSDFVLAGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR  
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

### **Signal sequence:**

amino acids 1-23

### **N-glycosylation site.**

amino acids 355-359

### **Casein kinase II phosphorylation site.**

amino acids 64-68, 142-146, 274-278

### **Tyrosine kinase phosphorylation site.**

amino acids 199-208

### **N-myristoylation site.**

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,  
305-311, 309-315, 320-326, 330-336

### **Cell attachment sequence.**

amino acids 149-152

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCCCACGGCGCCCGCGGGCTGGGGCGGTGCTTCTT  
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCGAAGG  
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCTCAACCTCCCAGGACCTATCTGG  
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG  
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACCTTTGGAG  
GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG  
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT  
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCAACAAGCAGCAGGAGGCCCCGGACCTCTTCC  
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGGCCCTCCTGC  
CTTCCCTGTCCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG  
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC  
AGTGTGGCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGT  
TTTGGCCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGCAATGCAAGAAGGG  
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT  
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG  
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCA  
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA  
ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG  
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC  
AGAAGACGAGTTGGTGGTGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA  
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG  
ACTGGCTACTGGTTGTGAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA  
ATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCCC  
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG  
CCTTACAGAGCAGCCCAGGTACCCAGGCCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGC  
CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTACAATGTGTGA  
ATTTCAAAGTTTTTCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTG  
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT  
TCTGTGTTACACACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA  
AAGGTCTTGGAAGTTAAAAA

## **FIGURE 40**

MAPWPPKGLVPAVLWGLSLFLNLP GPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER  
TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ  
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG  
YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPESNCLQCKKGWALHHLKCVDIDE  
CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE  
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDLVVLQQMFFG  
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

### **Signal sequence:**

amino acids 1-29

### **Transmembrane domain:**

amino acids 372-395

### **N-glycosylation site.**

amino acids 79-83, 205-209

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 290-294

### **Casein kinase II phosphorylation site.**

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

### **N-myristoylation site.**

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,  
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,  
326-332, 372-378, 395-401

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 321-333

### **EGF-like domain cysteine pattern signature.**

amino acids 181-193

## FIGURE 41

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA  
GCACCATGCAGCCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCCTGGCCAGCCCCGGG  
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT  
GCCCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGGCCAGT  
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTTCAGCCAGAGC  
TTCCGAGAGGTGGCCGGCAGGTTCTTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG  
CATGGAGCAGCGGCTGCCGCCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG  
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCCGGGCC  
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA  
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA  
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG  
GAGCATCTGGGCCCCGCTGGCGTCCGGCGCCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC  
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG  
CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG  
ATGTACATTGACCTGCAGGGGATGAAGTGGGCGGAGAACTGGGTGCTGGAGCCCCCGGGCTT  
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC  
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC  
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGGTGCAGAA  
GTGCAGCTGTGCCTCGGATGGTGCCTCGTGCCAAGGAGGCTCCAGCCATAGGCGCCTAGTG  
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTTCGAGGGTACCAGGAGAGCTG  
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT  
CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA  
GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT  
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCCANTGTGTGTCATTGTTTACTTGTCTGTGAC  
TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT  
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT  
AAAA

## **FIGURE 42**

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV  
ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP  
VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF  
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ  
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF  
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

### **Signal sequence:**

amino acids 1-18

### **N-glycosylation site.**

amino acids 158-162

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 76-80

### **Casein kinase II phosphorylation site.**

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

### **N-myristoylation site.**

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

### **Amidation site.**

amino acids 74-78

### **TGF-beta family signature.**

amino acids 282-298

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 43

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTGTCAGTGGCCTGATCGCGATGGGGGACAAA  
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG  
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT  
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA  
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG  
TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC  
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT  
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCCACCATTGGGAACCGGG  
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT  
GGGATAGTGATGCCTACGAATCCCAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT  
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCCTGTCAGCCTCTGATACTGGAGAATACA  
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT  
GTGGAGCGGAATGTGGGGGTGTCGTGGCAGCCGTCCTTGTAACCTGATTCTCCTGGGAAT  
CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA  
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG  
ACCTCGTCATTTCCTGGTGTGAGGCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT  
CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC  
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC  
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACCTGTTTAAA  
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC  
TTCTAAGTAGACAGCAAAAATGGCGGGGGTTCGAGGAATCTGCACTCAACTGCCCACCTGGC  
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC  
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG  
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG  
CCACTGGGATCCCTCTGCCCTGTCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT  
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG  
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAAGTGGAGGCTGGGCGCAGTGGCTCACGCCTG  
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA  
ACATGGAGAAACCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC  
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW  
KFDQGD TTRLVCYNNKITASYEDRVTF LPTGITFKSVTREDTGTYYTCMVSEEGGNSYGEVKV  
KLIVLVPPSKPTVNI PSSATIGNRAVLTCSEQDGSPSEYTWFKDGIVMPTNPKSTRAFSNS  
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVA AVLVTLI  
LLGILVFGIWFAYS RGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

### **Signal sequence:**

amino acids 1-27

### **Transmembrane domain:**

amino acids 238-255

### **N-glycosylation site.**

amino acids 185-189

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 270-274

### **Casein kinase II phosphorylation site.**

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,  
193-197, 203-207, 287-291

### **N-myristoylation site.**

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGA  
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCT  
GGAGGCCGCGCGAGCCCGCTTTCCACCCGACCTCTGCCCAGGCCGCGAGGCCCCAGCTCAG  
GCTCGTGCCCACCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCCTCACCTGG  
CGCTGCGACAGGGACTTGGA CTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC  
ATGTACCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCCTGCCCCCTGCACCGGCGTCA  
GTGACTGCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCA  
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA  
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG  
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTACCTCTCTCAGGAATGCC  
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCTCTGTGCGGAATGCCACATCCTC  
CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC  
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTTGTCTGCTCCGAGCCCAGGAGCGC  
CTCCGCCCCTGGGGTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC  
CTCGCTGCCCTTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA  
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGACACCAGCCCTCAGAGACCTGAGTTCTT  
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAAGTGGCCCTGGAGATTGAGGGTCCC  
TGGAACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGA ACTGAG  
GGGCTGGCCCCAGGCAGCTCCCAGGGGGTAGAACGGCCCTGTGCTTAAGAACTCCCTGCTG  
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

## **FIGURE 46**

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR  
TSGLCVPLTWRCRDLDCSDGSDEEEECRIEPTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL  
RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT  
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLL  
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSPL

### **Signal sequence:**

amino acids 1-30

### **Transmembrane domain:**

amino acids 230-246

### **N-glycosylation site.**

amino acids 126-130, 195-199, 213-217

### **Casein kinase II phosphorylation site.**

amino acids 84-88, 140-144, 161-165, 218-222

### **N-myristoylation site.**

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,  
224-230, 230-236, 263-269

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 44-55

### **Leucine zipper pattern.**

amino acids 17-39

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG  
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA  
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG  
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG  
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT  
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG  
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG  
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA  
TAATTCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA  
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG  
TACCCCGACCTACACAATATGGTTTTATTATGTCGCGATGATGGAACGTGGAATAATCTGCC  
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC  
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT  
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG  
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCTGCT  
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAAGTGTGGTGGAGTTTTACTGCGATCCT  
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTTCTTC  
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA  
CCACGTGGAAGATTGTGGCGTTACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC  
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCAGGGGGCCTCCCCGGAG  
TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG  
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG  
GGCTGCCCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA  
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA  
GTCTGTATTACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA  
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCCACTGGGT  
GTTGTTCCCTAAGAAACTGATTGATTAAAAAATTTCCCAAAGTGTCTGAAGTGTCTCTTCAA  
ATACATGTTGATCTGTGGAGTTGATTCTTTCTTCTTGGTTTTAGACAAATGTAAACAA  
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC  
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAAGNAAAAAA

## **FIGURE 48**

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI  
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI  
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS  
SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP  
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT  
WPSTHETLLTTWKIVAFTATSVLLVLLLVLARMFQTKFAHFPPRGPPRSSSSDPDFVVVD  
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGD TDTGPGESETCDS  
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWVFLRN

### **Signal sequence:**

amino acids 1-41

### **Transmembrane domain:**

amino acids 325-344

### **N-glycosylation site.**

amino acids 104-108, 134-138, 192-196

### **Casein kinase II phosphorylation site.**

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,  
364-368, 380-384, 467-471, 468-472

### **N-myristoylation site.**

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,  
478-484

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 394-405

## FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGA  
TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCCGAGCCCGCTCCACCGCCGT  
AGCGCCCGAGTGTCGGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGGAACCGCGCTACAGG  
CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTGCGCTGCTGAGTGCC  
TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA  
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT  
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAACTGATAGAA  
AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGA CTCTGGATTGGGCTCAGGAGGCGTGA  
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGA CTGATGGCAGCATAT  
CACAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATG  
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAATGATGA  
CCGGTGCAACATGAAGAACAATTTCA TTTGCAAATATTCTGATGAGAAACCAGCAGTTCCCTT  
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG  
GAAGAAGATGCCAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT  
CCTAATCCCCAGCATTCCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT  
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCTTAGCACAAAGAAGCAACACACCATC  
TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA  
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCA TTTCCGAGTGTGTT  
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA  
GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA  
GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG  
GTTATTAGGACATATAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC  
CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT  
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT  
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA  
GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG  
GCAGTGATAAAGATGGGCTGTGGAGCTTGGA AAACCACCTCTGTTTTCTTGCTCTATACAG  
CAGCACATATTATCATAACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAG  
GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGA  
GCAGGAAAAAAA

## **FIGURE 50**

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRRGGTQRPCYKVIYFHDTSRRL  
NFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGD F WIGLRRREEKQSNSTACQDL  
YAWTDGSISQFRN WYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY  
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAA LN LAYILIPSIPLLLLLLV  
VTTVVCWVWICRKRKREQPD PSTKKQHTIWPSPHQGN SP DLEVYNVIRKQSEADLAETRPDL  
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES  
GWVENEIYGY

### **Signal sequence:**

amino acids 1-21

### **Transmembrane domain:**

amino acids 235-254

### **N-glycosylation site.**

amino acids 117-121, 312-316

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 296-300

### **Casein kinase II phosphorylation site.**

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,  
299-303, 306-310, 323-327

### **N-myristoylation site.**

amino acids 18-24, 37-43, 76-82, 146-152

1000 900 800 700 600 500 400 300 200 100 0

## FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC  
ATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT  
GTTTGCCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT  
GTTCAGCATGCGCTTGTGGACCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC  
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC  
CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC  
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAATC  
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT  
CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT  
GCAGCAAATGTTTGCCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC  
TTTACCAACCTTCAACCCACAGGAGGTCTTTATTTCGTTCCACTAACATTTTTTCGGAATCTG  
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTCCAGTGTCAGAAAGAAGGACCCATCATCAT  
CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA  
GGCAGAGAACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG  
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT  
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG  
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG  
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC  
CATGGACTCTGCCACTGCCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG  
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTTGACCACAAATGGCCACCGTTTGCT  
GTTGACCTGACCATGGAACCTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA  
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCCGCTGGACATGT  
TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA  
ACTCAGGTGATGGAAGTTGGAAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT  
TTAAAATAAAGTGCCTTTATACAATG

## **FIGURE 52**

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMOVVFRHGAR  
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL  
TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFCQKE  
GPIIIHTDEADSEVLYPNYQSCWSLRQRTGRRRQTASLQPGISEDLKKVKDRMGIDSSDKVD  
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFHLHILES  
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW  
FVQLYYHGKEQVPRGCPDGLCPDLMFLNAMS VYTLSP EKYHALCSQTQVMEVGNEE

### **Signal sequence:**

amino acids 1-23

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 218-222

### **Casein kinase II phosphorylation site.**

amino acids 87-91, 104-108, 320-324

### **Tyrosine kinase phosphorylation site.**

amino acids 280-288

### **N-myristoylation site.**

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

### **Amidation site.**

amino acids 216-220

### **Leucine zipper pattern.**

amino acids 10-32

### **Histidine acid phosphatases phosphohistidine signature.**

amino acids 50-65

## FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT  
TAAATTTTCAGCTCATCACCTTCACCTGCCTTGGTCAATGGCTCTGCTATTCTCCTTGATCCTT  
GCCATTTGCACCAGACCTGGATTCTAGCGTCTCCATCTGGAGTGC GGCTGGTGGGGGGCCT  
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG  
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC  
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA  
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG  
ATTGTTACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA  
GTCCCAGAGGGTGTGAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA  
GCACCAGAACCAGTGGTATAACCGTGTGCCAGACAGGCTGGAGCCTCCGGGGCCGCAAAGGTGG  
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC  
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA  
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG  
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG  
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACCTGGGGAGAAAAGGAGGA  
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA  
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTGTTGCTCAGGGGAGGAG  
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA  
TGTGGCTGTCATCTGCTCAGTGAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA  
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT  
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCTGATTCTCAGGCCTTCAGAGTTGG  
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACCTACATCA  
CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT  
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACATA  
CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA  
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG  
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAGATAGCAC  
TATGTGTTCAAA

## **FIGURE 54**

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC  
RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC  
ENPESSFSVPVEGVRLADGPGHCKGRVEVKHQNQWYTVQCQTGWSLRAAKVVCRLGCGRAVL  
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG  
GDNLCSGRLEVLHKGWGSVCDNNGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL  
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

### **Signal sequence:**

amino acids 1-15

### **Casein kinase II phosphorylation site.**

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,  
267-271, 294-298, 316-320, 336-340

### **N-myristoylation site.**

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,  
180-186, 263-269, 286-292

### **Amidation site.**

amino acids 196-200

### **Speract receptor repeated domain signature.**

amino acids 29-67, 249-287

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC  
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG  
GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCTGCTGTTCCGGCTGCCTGGG  
CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG  
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT  
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA  
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC  
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC  
GACATACTTGTCACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA  
TGTGGACAAGAGGGTCATGGAGACAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC  
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG  
ATGAGCATTCCTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA  
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA  
TCCACACCAACCTCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC  
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG  
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCGAA  
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC  
AAGAACTCCTTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC  
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG  
AAAGACTGAAGAAACACATCTCGTGACAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA  
AGCTTCTTCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT  
AAAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG  
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

## **FIGURE 56**

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA  
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL  
VNNAGISYRGTIMDTTVDVVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAIISSIQKMSI  
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT  
AQGRSPVEVAQDVLAAVGGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

### **Signal sequence:**

amino acids 1-21

### **Transmembrane domain:**

amino acids 104-120, 278-292

### **N-glycosylation site.**

amino acids 228-232

### **Glycosaminoglycan attachment site.**

amino acids 47-51

### **Casein kinase II phosphorylation site.**

amino acids 135-139, 139-143, 253-257

### **Tyrosine kinase phosphorylation site.**

amino acids 145-153, 146-153

### **N-myristoylation site.**

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

### **Amidation site.**

amino acids 265-269

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 6-17

FIGURE 56

## FIGURE 57

CCCACGCGTCCGCTGGTGTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA  
AAAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC  
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCTTAA  
GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA  
GACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG  
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTTCATACCTTTGT  
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG  
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA  
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA  
GGCATTTCCTTCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG  
CTGGACATGTCTCGGTCCCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA  
TTTCATAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAACAACATG  
TCTGTGTCCTAATTTTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA  
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG  
ATTTTTATTCCATCTTCTATAGCTTTTTTAAACAACATTGGAAAGGATCCTTCCTGAGCGTTT  
CCTGGCAGTTTTTAAACGAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA  
AAGCGCAATAAGCACCTAGTTTTTCTGAAAAGTATTACCAGGTTTAGGTTGATGTCATCTA  
ATAGTGCCAGAATTTTAATGTTTGAAGTCTGTTTTTTCTAATTATCCCCATTTCTTCAATA  
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT  
ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA  
CCAAAATGACTTTATTAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA  
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACTTAAATTTTGT  
ATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAGAACTTCAAGCTCTCTAAATAAAA  
TGAAGGACTATATCTAGTGGTATTTTACAATGAATATCATGAACTCTCAATGGGTAGGTTTC  
ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT  
GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG  
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA  
AA  
AA

## **FIGURE 58**

MKFLLDI L L L L L P L L I V C S L E S F V K L F I P K R R K S V T G E I V L I T G A G H G I G R L T A Y E F A K L K S K  
L V L W D I N K H G L E E T A A K C K G L G A K V H T F V V D C S N R E D I Y S S A K K V K A E I G D V S I L V N N A G V V  
Y T S D L F A T Q D P Q I E K T F E V N V L A H F W T T K A F L P A M T K N N H G H I V T V A S A A G H V S V P F L L A Y C  
S S K F A A V G F H K T L T D E L A A L Q I T G V K T T C L C P N F V N T G F I K N P S T S L G P T L E P E E V V N R L M H  
G I L T E Q K M I F I P S S I A F L T T L E R I L P E R F L A V L K R K I S V K F D A V I G Y K M K A Q

**Signal sequence:**

amino acids 1-19

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 30-34, 283-287

**Casein kinase II phosphorylation site.**

amino acids 52-56, 95-99, 198-202, 267-271

**N-myristoylation site.**

amino acids 43-49, 72-78, 122-128, 210-216

## FIGURE 59

CCCACGCGTCCGCGGACGCGTGGGTCTGACTAGTTCTAGATCGCGAGCGGCCGCCGCGGGCTC  
AGGGAGGAGCACCGACTGCGCCGCACCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG  
TTTCGCTGGTCCTGTTGATGCCTGGCCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT  
GTTTCCATGCCACCTAAGGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC  
TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTTCGGCCCTTTCCCAGGACTGAACATGA  
AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC  
TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG  
AGGTTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCAACAAGTAACATGA  
CCTTGCGTGACAGAGACTTCCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA  
GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT  
AGCACGGGATTTATACAGTGCCTAATTAGTTTTTCCAGATATTTCTGAATATAAAAAATA  
ATGACTTTTTATGTCACTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTC  
ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA  
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT  
TGTTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACAACATC  
AGGAAGCAGAACTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC  
AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTTGCGGT  
GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAA  
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACCTATAGTTGAAAAGTACTTGCGAGA  
AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA  
TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC  
ATGGACTGGAAAGGATCCCAGGAATACAAGAAGGCAGAAAAAAAAGTTTGGAAGATCTTTAA  
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC  
GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA  
TTCATTTATGGAAAAGGATGGGATCCTTATGTTGGATAAACTACCTTCCCCAAAAGAGAACAT  
CAGAGGTTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAA  
TTATCTTTTTCATATCTGCAAGATTTTTTTTCATCAATAAAAAATTATCCTTGAAACAAGTGAGC  
TTTTGTTTTTGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA  
TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA  
AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGACGTTG  
TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT  
GTTTGGAAATATTATTGGATAAGAATAGCTCAATTATCCCAAATAAATGGATGAAGCTATAA  
TAGTTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA  
AAAATATTATATATAAAAGTAAAAA

## **FIGURE 60**

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGD SGQPLFLTPYIEAGKIQKGRELSL  
VGPFPGGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH  
GPYVVT SNMTLRDRDFPWTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALI QF  
FQIFPEYKN NDFYVTGESYAGKYVPAIAH LIHSLNPVREV KINLNGIAIGDGYSDPESI IGG  
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVTG  
CSNYYNFLRCTEPEDQLYYVKFSLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT  
EIMNNYKVLIIYNGQLDIIVAAAL TERSLMGMDWKGSQEYKKA EKKVWKIFKSDSEVAGYIRO  
AGDFHQVIIRGGGHILPYDQPLRAFD MINRFIYGKGWDPYVG

### **Signal sequence:**

amino acids 1-22

### **N-glycosylation site.**

amino acids 81-85, 132-136, 307-311, 346-350

### **Casein kinase II phosphorylation site.**

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,  
353-357, 424-428

### **Tyrosine kinase phosphorylation site.**

amino acids 423-432

### **N-myristoylation site.**

amino acids 22-28, 110-116, 156-162, 232-238

### **Serine carboxypeptidases, serine active site.**

amino acids 200-208

### **Crystallins beta and gamma 'Greek key' motif signature.**

amino acids 375-391

## FIGURE 61

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACAT  
TTTTCCCTTTCTAACAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT  
GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG  
CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG  
TCTAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC  
CTGGGGGAGGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT  
AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG  
TGAACCTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGGCCAGAAA  
TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA  
GACCAAAATAGATAACAGGATTCCCTGAACATTCCCTAAGAGGGAGAAAGTATGTTAAAAATA  
GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC  
AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTTTGGTCTGACCACTCTGCCTTGTGTTTT  
GCAGAATCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT  
CCTCACCGCCGCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCGCTGGAGGTGG  
ACAGCCGCTCTGTGGTCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCCCCAGCAGCCGGC  
ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT  
CCACCAAGGGACGGGGGCGGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA  
ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG  
CCCCTCATCGTGCAGCCCTGCAGCGAAGTGCTCACCTCACCAACAATGTCAACAAGCTGCT  
CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA  
AGCTGCTGCGGTGGATGACCTCTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC  
CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA  
TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA  
GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT  
TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT  
CTTCTACATCTACGGCTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCCGAGA  
CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG  
CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC  
CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC  
AGGCCTTCAATATCACAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG  
CAGTATCACCAACCCGCCGATGACTCTGCCCTGTGTGCCCTTCCCTATCCGGGCCATCAACTT  
GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC  
TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA  
CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC  
CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG  
TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC  
ATTACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG  
GCAACTTTATTTTCTTGGGGAACAAAGGTGAATGGGGAGGTAAGAAGGGGTTAATTTTGTG  
ACTTAGCTTCTAGCTACTTCCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA  
TTTCAATATTTCCCAAACCTTTAAGAAAAAATTTAAGAAGGTACATCTGCAAAAGCAA

## **FIGURE 62**

MGTLGQASLFAPPGNYFWSHDHSAFCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP  
RALEVDSRSVVLVSVVWVLLAPPAAGMPQFSTFHSENDRDWTFNHLTVHQGTGAVYVGAINRV  
YKLTGNLTIQVAHKTGPEEDNKSRYPPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLACGSL  
YQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY  
FPTLSSRKLPDPRESSAMLDYELHSDVSSLIKIPSDTLALVSHFDIFYIYGFASSGGFVYFL  
TVQPETPEGVAINSAAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP  
GDSLAAQAFNITSQDDVLFAIFSKGQKQYHHPDDSAFCFAPIRAINLQIKERLQSCYQGEEN  
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG  
YSVVFVGTKSGKLKKVRVYEFRCNSAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

### **Signal sequence:**

amino acids 1-32

### **Transmembrane domain:**

amino acids 71-87

### **N-glycosylation site.**

amino acids 130-134, 145-149, 217-221, 381-385

### **Casein kinase II phosphorylation site.**

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,  
384-388, 471-475, 481-485, 530-534

### **N-myristoylation site.**

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

## FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTGC GGACTGGAGTGGGAACCCGGGTCCCGCGCTTAGAGAACACGCGATGACCA  
CGTGGAGCCTCCGGCGGAGGCCGGCCCGCACGCTGGGACTCCTGCTGCTGGTCGTCTTGGGCTTCTGGTGCTCC  
GCAGGCTGGACTGGAGCACCTTGTCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA  
ACTTCATGCTGGAGGATTCACCTTCTGGATCTTGGGGGCTCCATCCACTATTTCCGTGTGCCAGGGAGTACT  
GGAGGGACCGCTGCTGAAGATGAAGGCCTGTGGCTTGAACACCTCACCACCTATGTTCCGTGGAACCTGCATG  
AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCCTTCGTCTGTATGGCCCGCAGAGATCG  
GGCTGTGGGTGATTCTGCGTCCAGGCCCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCAGCTGGCTAC  
TCCAAGACCCTGGCATGAGGCTGAGGACAACCTTACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACC  
TGATGTCCAGGGTGGTGCCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG  
GTTCTTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACCTGC  
TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT  
CAACACAGGAGTCGAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGTGGTGATGGAGT  
ACTGGACGGGGTGGTTTGACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT  
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG  
GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG  
ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG  
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC  
TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCTT  
TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCTCAGTGGCCACGTGCATGATCGGGGGCAGG  
TGTTTGTGAACACAGTATCCATAGGATTTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT  
ACACCGTGTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATGATGACCAGCGCAAAG  
GCTTAATTGGAATCTCTATCTGAATGATTCACCCCTGAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA  
GCTTCTTTTCAGAGTTGGCCTTGACAAATGGNGTTCCCTCCAGAAACACCCACATTACCTGCTTTCTTCTTGG  
GTAGCTTGTCCATCAGTCCACGCCCTTGTGACACCTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTGTATTCA  
TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCAGAAGACGCTTACCTCCAGGTCCCTGGTTGA  
GCAGCGGAATCAACCAGGTCACTGTTTTTGGAGGAGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCC  
ACCTGGGCAGGAACCAGTACATTAAGTGAGCGGTGGCACCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC  
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAGCATCTCCTTAAGTAGCAACCTCAGGG  
ACTGGGGGCTACAGTCTGCCCTGTCTCAGCTCAAAACCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCC  
TGGCTTTGTTGATGATGGCTTTTCTACAGCCCTGCTCTTGTGCCGAGGCTGTGCGGCTGTCTTAGGGTGGGAGC  
AGCTAATCAGATCGCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGCACCGGACGTCACAGCCC  
TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT  
TTTATCCCCGAAATCCTGGGTGTGTACCAAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT  
CTTCTTCAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACCTCGGCGTGAGAAACATGTGACTTCCCCTT  
TCCCTTCCCACTCGCTGCTTCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCTGCGTCTTCC  
CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCA  
CATCCAGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCC  
AGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA  
GGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG  
ACAGAAGGCCAGCTCAGTGGCCCCGCTCCCAACCCCCACGCCCCAAGCAGAGGGGAGAGCAGCCCTCCTTC  
GAAGTGTGTCCAAGTCCGCATTTGAGCCTTGTCTGGGGCCAGCCCAACACCTGGCTTGGGCTCACTGTCCTGA  
GTTGCAGTAAAGCTATAACCTTGAATCAAA

## **FIGURE 64**

MTTWSLRRRPPARTLGLLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW  
IFGGSIIHYFRVPREYWRDRLLKMKACGLNTLTITYVPWNLHEPERGKFDFSGNLDLEAFVLMA  
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ  
YKRGGP I IAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLTSDNKDGLSKGIVQGVLAT  
INLQSTHELQLLTTFLFNVQGTQPKMVM EYWTGWFD SWGGPHNILDSSEVLKTVSAIVDAGS  
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDY TAKYMKLRDFFGSISGIP  
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE  
TSITSSGILSGHVHDRGQVFVNTV SIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN  
IDDQRKGLIGNLYLNDSP LKNFRIYS LDMKKSFFQRFGLDKWXS LPETPTLP AFFLGSL SIS  
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPW LSSGINQVIVFEETMAGPA  
LQFTETPHLGRNQYIK

### **Signal sequence:**

amino acids 1-27

### **Casein kinase II phosphorylation site.**

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

### **N-myristoylation site.**

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,  
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

## FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC  
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG  
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC  
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCC  
AAGAAGCTGTCCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCA  
GGCAGACACTCGGTCTGTTCTAGTGGATAGGGGTTCATGACCGGTTTCTCCTAGACGGGGCCC  
CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC  
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACCTA  
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA  
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG  
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCCTC  
AGATCCAGACTTCCTTGCCGAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC  
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTTCAGGTGGAGAATGAATATGGTAGC  
TACAGAGCCTGTGACTTCAGCTACATGAGGCACCTGGCTGGGCTCTTCCGTGCACTGCTAGG  
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG  
GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT  
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA  
CTGGGGCCAGAATCACTCCACACGGTCTGTGTGCTGTAACCAAAGGACTAGAGAACATGC  
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG  
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC  
TATATCTGAAGCAGGGGACCCACACCTAAGCTTTTGTCTCTTCGAGATGTCATCAGCAAGT  
TCCAGGAAGTTCCCTTTGGGACCTTTACCTCCCCCGAGCCCCAAGATGATGCTTGGACCTGTG  
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTTCTAGACTTGCTTTGCCCCCGTGGGCCCAT  
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC  
GAACCTATATGACCATAACATTTTGGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC  
CATGACCGTGCCATATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG  
AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAACTGGATATCTTGGTGGAGAACATGG  
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG  
GGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG  
GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT  
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG  
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCGGTACTGGACAAAGCAGGGGCC  
ACAACAGACCCTCTACGTGCCAAGATTCCCTGCTGTTTCCCTAGGGGAGCCCTCAACAAAATTA  
CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC  
CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACACTGAG  
TGCCTCTGAACCAATGGAGTTAAGTGGGCACTGAAAGGTAGGCCGGGCATGGTGGCTCATGC  
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAGGACTTCAAGA  
CCAGCCTGGCCAACATGGTGAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG  
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC  
AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA  
GCACTCCATCTCAAAAAAAAAAAAA

## **FIGURE 66**

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDGRGHDRFLLDGAPFRYVSGSLHYFRVPRVL  
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNNGSRDLIAFLNEAALANLLVILRPGPYI  
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE  
YGSYRACDFS YMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF  
TL LRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNM MYFHGGTNF  
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMML  
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN  
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP  
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGD TFLYL  
PGWTKGQVWINGFNLGRYWTQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD  
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

### **Signal sequence:**

amino acids 1-27

### **N-glycosylation site.**

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 4-8

### **Casein kinase II phosphorylation site.**

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,  
603-607, 644-648

### **Tyrosine kinase phosphorylation site.**

amino acids 191-198

### **N-myristoylation site.**

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,  
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

## FIGURE 67

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC  
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT  
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT  
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT  
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT  
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG  
AAAAACTCAGGCAGCACATTTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTCATG  
CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC  
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC  
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA  
TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA  
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG  
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT  
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTACATAATGACGGCAC  
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACTCCAGA  
ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT  
TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT  
GACTTGTTTTAAAATTATGGCATAACAAAATTGTTACTATTCTCCCTCTATTACCCATGTCA  
AAAACCTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT  
AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA  
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC  
CAAAACAATTGTTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC  
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAACTG  
CTTGACCGCCTGCCAGCCAGCTGGGCCAGTGTGCGATGCTCAAGAAAAGCGGGCTTGTTG  
TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA  
AATATTCCCTTTGCAAATGGGATTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC  
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG  
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT  
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTTCTTTTGGGG  
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTTAAATTGTTTGTAACCTGGAT  
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCTGCTAAAGTAAATGATTAAATTGACATT  
TTCTTACTAAAAAAAAAAAAAAAAAAAA

## **FIGURE 68**

MAYMLKKLLISYISIIICVYGFICLYTLFWLFRIP LKEYSF EKVREESSFS DIPDVKNDF AFL  
LHMVDQYDQLYSKRFGVFLSEVSENKLREISLNHEWTFEKL RQHISRNAQDKQELH LFM LSG  
VPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH  
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLES LREL RHLKILHVKS NLTKVPSN  
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAE LELQNC ELERIPHAIFSLSNLQELDLKS  
NNIRTIEEIIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLES LYFSNNKLES LPVAVFSLO  
KLRC LDVS YNNISMIP IEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITS LP  
EKVGQLS QLTQLELKGNC LDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP  
FANGI

### **Signal sequence:**

amino acids 1-20

### **N-glycosylation site.**

amino acids 241-245, 248-252, 383-387

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 326-330

### **Casein kinase II phosphorylation site.**

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

### **Tyrosine kinase phosphorylation site.**

amino acids 349-355, 375-381

### **N-myristoylation site.**

amino acids 78-84, 124-130, 212-218, 392-398

## FIGURE 69

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCCATTCCCTTTTCATTGACAAACTGACTTTTTTTTATTTCT  
TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTACACACATAAGGAT  
CTGTGTTTGGGGTTTCTTCTCTCCCTGACATTGGCATTGCTTAGTGTTGTGTGGGGAGGGAGACCACGTGG  
GCTCAGTGCTTGGCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC  
ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCTGTCTTTACTTCAAATAACACAAC  
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG  
AACAGCCAGGCCAAAACATTGCCACGGAGTCTTGCTCCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC  
AGTTTTGATTCCCTGCCACCTTGCTGTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA  
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG  
CTCCCATCAGTTTTCATGGAATAAATCAGTGCCCTGCTGGGAACCACTGCTGGAGATCCCTACAGAGAGCTTC  
CACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCTCACTGTGGGGAATGCTGATAAACCAAGTCA  
CACAGCTGCTCTATTCTACACAAATCTACCCCTTGGCTGGCTGGAACCTGACGTTTCCCTGGAGGTGTCCAGAAA  
GCTGATGTAACACAGAGCCTATAAAAGCTGTCCGTCTTAAGGCTGCCAGCGCCTTGCCAAAATGGAGCTTGTA  
AGAAGGCTCATGCCATTGACCCTCTTAATTCTCTCCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT  
GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC  
AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAATTATCTTTCTATGTCCAGCTT  
GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG  
CAAGTCTGCAGTAAAAACGACTATGTTCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT  
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCCTAACATCTCTATTCCAACTGT  
GGCGGTTACCTGGATACCTTGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT  
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAACTAAACTTCAAAGAGATTTTCCCTAGAAATAGAC  
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT  
GGCCGTGTGACTCCACCTTCGAATCGTCATCAAACCTCTCTGACTGTCTGTGTTGTCTACAGATTATGCCAATTCT  
TACCGGGGATTTTCTGCTTCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT  
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA  
GACCCAACTTGACAGACCAAATATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA  
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC  
CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA  
GAAGATGATGTAATACAAAGTCAAAATGCACTGGGCAAAATATAACACCAGCATGGCTCTTTTGAATCCAATTCA  
TTTGAAAAGACTATACTTGAATCACCATAATTATGTGGATTGTAACCAAACTCTTTTGTTCAGTTAGTCTGCAC  
ACCTCAGATCCAAATTTGGTGGTGTCTTGTATACCTGTAGAGCCTCTCCACCTCTGACTTTGCATCTCCAACC  
TACGACCTAATCAAGAGTGGATGTAGTGCAGATGAACTTGTAAAGGTGTATCCCTTATTTGGACACTATGGGAGA  
TTCCAGTTTAAATGCCTTTAAATTTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAGTTTTGTATGTGAT  
AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG  
AAAACAGATTCCATCATAGGACCCATTTCGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTTCAGCAT  
GAAACACATGCGGAAGAACTCCAAACAGCCCTTCAACAGTGTGCATCTGTTTCTTTCATGGTTCTAGCTCTG  
AATGTGGTGACTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG  
CAGAACTATTAACTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT  
GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

## **FIGURE 70**

MELVRRRLMPLTLLILSCLAELTMAEAEAGNASCTVSLGGANMAETHKAMILQLNPSENCTWTIERPENKSIRIIFSIVQLDPDGSCSENIKVFDGTSSNGPLLQVCSKNDYVPVFESSSSTLT  
FQIVTDSARIQRTVFVFFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV  
EKDYKIKLNFKEIFLEIDKQCKFDLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLS  
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIIISKSYLEAFNSNGNNLQLKDPTCRP  
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQQLQIIVKCEMGHNST  
VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN  
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL  
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDIIIGPIRLKRDRSASGNSGFQHETHA  
EETPNQPFNSVHLFSFMVLALNVVTVAITITVRHFVNQRADYKYQKLQNY

### **Signal sequence:**

amino acids 1-24

### **Transmembrane domain:**

amino acids 571-586

### **N-glycosylation site.**

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,  
394-398, 419-423

### **Casein kinase II phosphorylation site.**

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,  
408-412, 463-467, 520-524, 556-560

### **Tyrosine kinase phosphorylation site.**

amino acids 172-180, 407-415, 407-416, 519-528

### **N-myristoylation site.**

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

Station	Time	Lat.	Long.	Alt.	Temp.	Wind	Clouds	Remarks
1	0800	34° 15' N	121° 05' E	10	18.5	10	100	Clear
2	0900	34° 30' N	121° 15' E	10	19.0	10	100	Clear
3	1000	34° 45' N	121° 25' E	10	19.5	10	100	Clear
4	1100	35° 00' N	121° 35' E	10	20.0	10	100	Clear
5	1200	35° 15' N	121° 45' E	10	20.5	10	100	Clear
6	1300	35° 30' N	121° 55' E	10	21.0	10	100	Clear
7	1400	35° 45' N	122° 05' E	10	21.5	10	100	Clear
8	1500	36° 00' N	122° 15' E	10	22.0	10	100	Clear
9	1600	36° 15' N	122° 25' E	10	22.5	10	100	Clear
10	1700	36° 30' N	122° 35' E	10	23.0	10	100	Clear
11	1800	36° 45' N	122° 45' E	10	23.5	10	100	Clear
12	1900	37° 00' N	122° 55' E	10	24.0	10	100	Clear
13	2000	37° 15' N	123° 05' E	10	24.5	10	100	Clear
14	2100	37° 30' N	123° 15' E	10	25.0	10	100	Clear
15	2200	37° 45' N	123° 25' E	10	25.5	10	100	Clear
16	2300	38° 00' N	123° 35' E	10	26.0	10	100	Clear
17	0000	38° 15' N	123° 45' E	10	26.5	10	100	Clear
18	0100	38° 30' N	123° 55' E	10	27.0	10	100	Clear
19	0200	38° 45' N	124° 05' E	10	27.5	10	100	Clear
20	0300	39° 00' N	124° 15' E	10	28.0	10	100	Clear
21	0400	39° 15' N	124° 25' E	10	28.5	10	100	Clear
22	0500	39° 30' N	124° 35' E	10	29.0	10	100	Clear
23	0600	39° 45' N	124° 45' E	10	29.5	10	100	Clear
24	0700	40° 00' N	124° 55' E	10	30.0	10	100	Clear

GACGGAAGGAACAGACGCTCCCGAGGCCGCGGAGCGCTGCAGAGAGGACGCCCGCCTGCGCCG  
GGACATGCGGCCCCAGGAGCTCCCAGGCTCGCGTTCCCGTTGCTGTGCTGTTGCTGTGCTG  
TGCTGCCGCCGCCGCGCTGCCCTGCCACAGCGCCACGCGCTTCGACCCCACTGGGAGTCC  
CTGGACGCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG  
GGGAGTGTTTTCCGTGCCCAGCTTCGGTAGCGAGTGTTCTGGTGGTATTGGCAAAGGAAA  
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCTCCTAGTTTCAAATATGAAGAT  
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAAGTGGGCAGATATTTTTTCAGGC  
CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT  
CAGAATATTCGTGGAAGTGAATGCCATAGATGAGGGGCCAAGAGGGACATTGTCAAGGAA  
CTTGAGGTAGCCATTAGGAACAGAAGTACCTGCGTTTTGGACTGTACTATTCCCTTTTTTGA  
ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG  
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG  
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT  
ATATAATGAAAGCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA  
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA  
CATAAATGGGAAAAGTGCATGACAATAGACAACTGTCTTGGGGCTATAGGAGGGAAGCTGG  
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG  
GAGGAAATCTTTTGATGAATATTGGGCCCCACACTAGATGGCACCATTCTGTAGTTTTTTGAG  
GAGCGACTGAGGCAAGTGGGGTCTTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA  
TACCTGGCGATCCCAGAATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG  
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCAATCAGGACAGCTGTTCTTGGCCAT  
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG  
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTTCATCAGATGC  
CGTGTAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG  
ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA  
AAGCAATGTAAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTCCCTTTTTTCCCACTA  
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC  
TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTTCACATTATAGTAG  
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG  
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT  
TTTTTTTTTGTGCCAACATCATAGAGTGATTTACAAAATCCTAGATGGCATAGCCTACTACA  
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG  
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG  
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG  
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA  
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA  
GTTTTTCTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTTAATT  
TTTTAAACCTTTTTGGCTCTTTTGTAAATAACACTTAGCTTAAACATAAACTCATTGTGCAA  
ATGTAA

## **FIGURE 72**

MRPQELPRLAFPLLLLLLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG  
VFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPSPFKYEDFGPLFTAKFFNANQWADIFQAS  
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLEW  
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY  
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI  
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT  
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI  
SLEQNGIMVELPQLTIHQMPCKKGWALALTNVI

### **Signal sequence:**

amino acids 1-28

### **N-glycosylation site.**

amino acids 171-175, 239-243, 377-381

### **Casein kinase II phosphorylation site.**

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,  
375-375

### **Tyrosine kinase phosphorylation site.**

amino acids 361-369, 389-397

### **N-myristoylation site.**

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

### **Leucine zipper pattern.**

amino acids 410-432

### **Alpha-L-fucosidase putative active site.**

amino acids 283-295

## FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT  
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC  
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTG  
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT  
TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC  
AGCTGAATTTACAGAAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG  
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGGAGA  
TGGATTTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG  
TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGAT  
ACTTGGAATAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA  
AACTGCAACACAAACAACAGAATTTATTGTGTCAGTGACAGTACCTACTCGGTGGCATCCCCCTT  
ACTCTACAATACCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG  
AGAAAAAATTGATTTGTGTACAGAAAGTTTTTATGGAACTAGCACCATGTCTACAGAAAC  
TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC  
CCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGC  
TATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT  
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA  
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA  
GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTTCATGCTCC  
TTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTT  
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC  
CCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAGCTTCCCACGGCC  
TTTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTTTGCAAAGTGCAA  
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG  
TGGGTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGAC  
CCTTCTTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA  
AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG  
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT  
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA  
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT  
AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTTATCTGA  
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA  
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT  
TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTTCA  
GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTTATTTTTTGCTGAGACTAATCTT  
ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG  
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA  
GCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA  
AGCATTTAGAAAACCTT

## **FIGURE 74**

MARCFSLVLLLLTSIWTTLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTAKEACR  
LLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLWKVPVSRQF  
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP  
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFF  
GAAAGLGFCYVKRYVKAFPFTNKNQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP  
SKTTVRCLEAEV

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 235-254

### **N-glycosylation site.**

amino acids 53-57, 130-134, 289-293

### **Casein kinase II phosphorylation site.**

amino acids 145-149, 214-218

### **Tyrosine kinase phosphorylation site.**

amino acids 79-88

### **N-myristoylation site.**

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

## FIGURE 75

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGATTTCGGTGCCGCGACTTTCACGATGG  
CTCGCCCAACCTTACTACCTTCTGTGCGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA  
ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG  
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA  
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT  
TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC  
TGATGACGTGCAAACCCCCCTATATATGGGCCCCTGAGTATATCAAGTACTTCAATGATAAA  
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA  
TTGGTCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAACCT  
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC  
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA  
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG  
AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAACTATCA  
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTGAGA  
TGGGGAAAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATT  
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG  
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG  
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAACCTGTTTCACTGGAGCAAGAAAGAGATC  
TCATAGGACGGAGGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC  
AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG  
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA  
CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC  
TTCATTAAGATATAAGCCTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG  
TTTTTATCATTGAAGACAATATTGAACAACCCCCCTATTTTGTGGGGATTGAGAAGGGGTGAA  
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC  
TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCGTAATAAAAAGATTGGGATT  
TCCTTTTG

## **FIGURE 76**

MAVLAPLIALVYSVPRLSRWLAQPYLLSALLSAAFLLVKLPPLCHGLPTQREDGNPCDFD  
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL  
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC  
TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE  
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

### **Signal sequence:**

amino acids 1-48

### **Transmembrane domain:**

amino acids 111-125

### **N-glycosylation site.**

amino acids 165-169, 185-189

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 154-158, 265-269

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

### **N-myristoylation site.**

amino acids 188-194, 225-231

### **Myb DNA-binding domain repeat signature 1.**

amino acids 244-253

## **FIGURE 77**

GGACAGCTCGCGGCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG  
GGGCCCCAGCCTGGCCCGGGTCACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA  
TTGCTCCTGCTGCCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTCCAACAG  
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG  
TGGTGGAGACACCCGAGGAGACCCTGTTACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC  
CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT  
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT  
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG  
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA  
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTTCTTACCAGTCCCCCA  
ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG  
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG  
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCC  
CAGGCCTGGCACCTGGCGTGCGAAGCTACGGCCCCCGCCACCGCCGCTGCACCGCTATGAT  
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCTGAGAAGCT  
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC  
AGCTCTTTGCCGCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT  
GGCAGCGTCCGCTACCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT  
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC  
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT  
CGTTTTCCCTTGTGGGTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT  
TTAAACATTTTTTTTACTATTTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG  
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTGTGGGCTTCTG  
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGGCCAGAGTGGGC  
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCT  
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC  
CACCTCTACTTCTCTGTGAAGCCGCTGACCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA  
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAACTCCCCTCCCCGTTCCCCT  
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTGTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG  
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGTCTATGACTGCCTTCCGCCAA  
AA  
AA

## **FIGURE 78**

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNHGHGKDLLNGVKLVVETPEETLFTYQ  
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD  
KEHDVSLEIQDLRLLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ  
VCAEQAAVVASFQELFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR  
HRRLLHRYDVFCFATALKGRVYYLEHPEKLTLTAREACQEDDATIAKVGQLFAAWKFHGLDR  
CDAGWLADGSVRYPVVHHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

**Signal sequence:**

amino acids 1-17

**Casein kinase II phosphorylation site.**

amino acids 29-33, 53-57, 111-115, 278-282

**Tyrosine kinase phosphorylation site.**

amino acids 137-145

**N-myristoylation site.**

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**THE UNIVERSITY OF CHICAGO**

[illegible]

## **FIGURE 80**

MMWRPSVLLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDA PHDDAHGNFQYDHEAFLGR  
EVAKEFDQLTP EESQARLGRI VDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT  
YDTRDGRV GWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATRE  
ELTAFLHP EEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEP AWVQTERQQ  
FRDFRDLNKDGHLDGSEVGHVLP PAQDQPLVEANHLLHESD TDKDGRLSKAEILGNWNMFV  
GSQATNYGEDLTRHHDEL

### **Signal sequence:**

amino acids 1-20

### **N-glycosylation site.**

amino acids 140-144

### **Casein kinase II phosphorylation site.**

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,  
291-295, 298-302

### **N-myristoylation site.**

amino acids 263-269, 311, 317

### **Endoplasmic reticulum targeting sequence.**

amino acids 325-330

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

## FIGURE 82

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPPGGCPLEEF SVY  
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASF TVT  
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDKADIAFLIDGSFNIGQRRFNLQKN  
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLFAIKEVGFRGGNSNTGKAL  
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG  
MVQDVT FVDKAVCRNNGFFSYHMPNWF GTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI  
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI  
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITI  
FSVGVAWAPLDDLKDMASKPKESHAF FTREFTGLEPIVSDVIRGICRDFLESQQ

### **Signal sequence:**

amino acids 1-24

### **N-glycosylation site.**

amino acids 100-104, 221-225

### **Casein kinase II phosphorylation site.**

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,  
425-429, 478-482, 528-532

### **N-myristoylation site.**

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,  
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

### **Amidation site.**

amino acids 145-149

## FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCC  
GGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGGCCCGCAGCGCAACTCGGTCCAGTCG  
GGGCGGCGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC  
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTC  
AAGCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA  
GGTTGAGGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTAT  
CACAATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAAT  
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG  
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC  
ATGTACTGCCAGTTTGGCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT  
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA  
TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG  
TGCTGTGCCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA  
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG  
GAGCCTTGAGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG  
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC  
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG  
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGAGCCTGCGGCTGCCGCCGCT  
GCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA  
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA  
TCTTCTTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGC  
TCCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAAC  
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG  
ACAGCCGTTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGG  
AGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTGGCCTGCTTTGCAAAATCAA  
CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG  
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTTCATCC  
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC  
CCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG  
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT  
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACCAGCCTTGGTGCCACCAA  
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG  
GTCAAATAATTCTCACATCCCTCTAAAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCAC  
AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT  
TGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCATAACAGTTAACCTGCAGAAACA  
GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC  
TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT  
GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTTCAGGTGTCA  
TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA  
TAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC  
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 84**

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ  
HKLRS AVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG  
QMV FSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRM LCTR DSECCG  
DQLCVWGHCTKMATRGSNGTICDNQRDCQ PGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL  
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV  
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEI

### **Signal sequence:**

amino acids 1-19

### **N-glycosylation site.**

amino acids 96-100, 106-110, 121-125, 204-208

### **Casein kinase II phosphorylation site.**

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,  
327-331

### **N-myristoylation site.**

amino acids 202-208, 217-223

### **Amidation site.**

amino acids 140-144

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACATCTCTCTGGGCTCAGAAGGACTCTG  
AAGATAACAATAATTTTCAGCCCATCCACTCTCTCTTCCCTCCCAAACACACATGTGCATGTACACACACACATACA  
CACACATACACCTTCTCTCTTCACTGAAGACTCACAGTCACTCACTCTGTGAGCAGGTCATAGAAAAGGACAC  
TAAAGCCTTAAGGACAGGCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTGAGTCAAAAAACATGGGAGGGG  
CCAGGCACGGTGACTCACACCTGTAATCCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTCAGGAG  
TTCGAGACCAGCCTGGCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC  
AGGTGCCTGTAATCCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGT  
CAGCTGAGTGACCCGCTGCACCTCCAGCCTGGGTGACAGAATGAGACTCTGTCTCAAACAAACAAACACGGGAGGA  
GGGGTAGATACTGCTTCTCTGCAACCTCCTTAACCTCTGCATCCTCTTCTTCCAGGGCTGCCCTGATGGGGCCTG  
GCAATGACTGAGCAGGCCCCAGAGGACAAGGAAGAGAAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG  
GTGTAGAATGACTGCCCTGGGAGGGTGGTTCTTGGGCCCTGGCAGGGTTGCTGACCTTACCCTGCAAAACACA  
AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCCTGCCCTGCAGCTCCACCATGAGGCTTCTCGTGGCCCC  
ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCTCA  
GTGTGCCTGCCAGATCCGGCCCTGGTATACGCCCGCTCGTCTACCGCGAGGCTACCCTGTGGACTGCAATGA  
CCTATTCTGACGGCAGTCCCCCGGCCTCCCCGAGGCACACAGACCCTGCTCCTGCAGAGCAACAGCATTGT  
CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAACTCTCACAGAGCTGGACCTGTCCAGAACAGCTTTTCGGA  
TGCCCGAGACTGTGATTTCCATGCCCTGCCCGAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT  
GGAGGACCACAGCTTGTGAGGGCTGGCCAGCCTACAGGAACCTCTATCTCAACCACAACCAGCTCTACCGCATCGC  
CCCCAGGGCCTTTTCTGGCCTCAGCAACTTGCTGCGGCTGCACCTCAACTCCAACCTCCTGAGGGCCATTGACAG  
CCGCTGGTTTTGAAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGCCATCCTGGACAT  
GAACCTCCGGCCCCCTGGCCAACCTGCGTAGCCTGGTGTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC  
CCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGCTGGCCCGGGTGCCAGGCGGGCACT  
GGAACAGGTGCCCCGGGCTCAAGTTCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGGACTTTGC  
CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTGGT  
GAACCTCCCCGAGCTGACCAAGCTGGACATACCAATAACCCACGGCTGTCTTTCATCCACCCCCGCGCTTCCA  
CCACCTGCCCCAGATGGAGACCCTCATGCTCAACAACAACGCTCTCAGTGCCTTGCAACCAGCAGACGGTGGAGTC  
CCTGCCCAACCTGCAGGAGGTAGGTCTCCACGGCAACCCCATCCGCTGTGACTGTGTCTATCCGCTGGGCCAATGC  
CACGGGCACCCCTCCGCTTCATCGAGCCGCAATCCACCCTGTGTGCGGAGCCTCCGGACCTCCAGCGCCTCCC  
GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGGCCCTCATCTCCCCACGAAGCTTCCCCCAAG  
CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGTGCTGATTGCCGGGCACTGGCCGAACCCGAACCCGAGATCTACTG  
GGTCACTCCAGCTGGGCTTCGACTGACACCTGCCCATGCAGGCAGGAGGTACCGGGTGTACCCCGAGGGGACCTT  
GGAGCTGCGGAGGGTGACAGCAGAAGAGGCAGGGCTATACACCTGTGTGGCCCAAGAACCTGGTGGGGGCTGACAC  
TAAGACGGTTAGTGTGGTTGTGGGCCGTGCTCTCCTCCAGCCAGGCAGGGACGAAGGACAGGGGCTGGAGCTCCG  
GGTGCAGGAGACCCACCCCTATCACATCCTGTCTTGGGTACCCCAACCAACAGTGTCCACCAACCTCAC  
CTGGTCCAGTGCCTCCTCCTCCGGGGCCAGGGGGCCACAGCTCTGGCCCGCCTGCCTCGGGGAACCCACAGCTA  
CAACATTACCCGCTCCTTACGGCCACGGAGTACTGGGCCCTGCCTGCAAGTGGCCTTTGCTGATGCCACACCCA  
GTTGGCTTGTGTATGGGCCAGGACCAAGAGGGCACTTCTTGCCACAGAGCCTTAGGGGATCGTCTCTGGGCTCAT  
TGCCATCCTGGCTCTCGCTGTCTTCTCCTGGCAGCTGGGCTAGCGGGCCACCTTGGCACAGGCCAACCCAGGAA  
GGGTGTGGGTGGGAGGCGGCTCTCCTCCAGCCTGGGCTTTCTGGGGCTGGAGTGCCCCCTTCTGTCCGGGTGT  
GTCTGTCCCCCTCGTCTGCCCTGGAATCCAGGGAGGAAGCTGCCCAGATCCTCAGAAGGGGAGACACTGTTGCC  
ACCATTTGTCTCAAAATTCTTGAAGCTCAGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTACCAAA  
AGAGAAGCAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCTGGCAGCTGGGC  
CAAGACAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGAAAAAGTTGCCCTTACCTCCTAGGGTCA  
CCTCTGCTGCCATTCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCCTCCCCATCTT  
CTCTCTGCCAGAGGCTCCTGGGCCTGGCTTGGCTGTCCCCCTACCTGTGTCCCCGGGCTGCACCCCTTCTCTTC  
TCTTCTCTGTACAGTCTCAGTTGCTTGTCTTGTGCTCCTGGGCAAGGGCTGAAGGAGGCCACTCCATCTCAC  
CTCGGGGGGCTGCCCTCAATGTGGGAGTGACCCAGCCAGATCTGAAGGACATTTGGGAGAGGGATGCCAGGAA  
CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGGAGAA  
ATGTGTCACTCCCCCAACCCGATTCACTCTTTTCTCCTGTTTTGTAAAAAATAAAAAATAATAATAATAATAA  
AAAA

1000 900 800 700 600 500 400 300 200 100 0

## **FIGURE 86**

MRLLVAPLLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA  
VPPALPAGTQTLILLQSNSIVRVDQSELGYLANLTDLSONSFSDARDCDFHALPQLLSLHL  
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE  
MLPNLEILMIGGNKVDAILD MNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ  
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELV SIDKFALVNLP  
ELTKLDITNNPRLSFIHPRAFHHL PQMETLMLNNNALSALHQQTVESLPNLQEVGLHG NPIR  
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPFREM TDHCLPLISPRSFP PSLQ  
VASGESMVLHCRALAEPEPEIYWVTPAGLR LTPAHAGRRYRVYPEGTLELRRVTAE EAGLYT  
CVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTP PNTVSTNLTW  
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS  
CHRALGDRPGLIAILALAVLLLAAGLAAHLGTGQPRKGVGGRRPLPPAWAFWGSAPS VRVW  
SAPLVLPWNPGRKLPRSSEGETLLPPLSQNS

### **Signal sequence:**

amino acids 1-18

### **Transmembrane domain:**

amino acids 629-648

### **N-glycosylation site.**

amino acids 94-98, 381-385, 555-559, 583-587

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 485-489

### **Casein kinase II phosphorylation site.**

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,  
243-247, 313-317, 488-492, 700-704

### **Tyrosine kinase phosphorylation site.**

amino acids 532-540

### **N-myristoylation site.**

amino acids 15-21, 493-499, 566-572

### **Amidation site.**

amino acids 470-474, 660-664, 692-696

## FIGURE 87

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT  
CTACATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA  
CATCAAGTTGCGAGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCCACCC  
CCTGGCCACACTCTTCAAGATCCTGGCGTCTTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA  
CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG  
CGACATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA  
GCGCTTCGCCGCTCTTCTGTGCGGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT  
GGACAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTTCATGCTCAGTGGCAT  
CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGACGTGACCATCCCCGC  
CAGCATTGCCCCAGCTCACGGGCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCTGCGCT  
GGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA  
TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACACCGCTACATCGTCATCGA  
CGGGCTGCGGGAGCTCAAACGCCTCAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTTCAC  
AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCTCAACAGCCTCAA  
GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAGCCT  
CCACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATCAGCTTCCAGCACCT  
GCACCGCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAA  
CCTGGAGCGCCTCTACCTGAACCGCAACAAGATCGAGAAGATCCCCACCCAGCTCTTCTACTGCCGCAAGCTGCG  
CTACCTGGACCTCAGCCACAACAACCTGACCTTCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT  
AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT  
GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG  
CAACCGGCTGGAGTGCCTGCCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA  
GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGCCCTGAGCGGAG  
GCCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCTCAGGCCCGGAGGGGAGGCCCTAGCTTCTCCCAG  
AACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCGAGGAGCTGGGGCCGCTTGTGAGTCAGGCGCAGAGCGAGA  
GGACAGTATCTGTGGGGCTGGCCCCCTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGCTTGTGGAGGAG  
AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG  
CTGCCACCAGAGGTCTGGGACCTCACTTTAGTTCTTGGTATTTATTTTTCTCCATCTCCACCTCCTTCATCC  
AGATAACTTATACATTCCCAAGAAAGTTCAGCCCAGATGGAAGGTGTTTCAGGGAAAGGTGGGCTGCCTTTTCCCC  
TTGTCTTTATTTAGCGATGCCGCCGGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGGCGAACCCAG  
CCATGGGACGGTCAACCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCCCTCCAGCTGGA  
AAGGCCAGGCCTGGAGCTTGCCCTTTTCAGTTTTTGTGGCAGTTTTAGTTTTTTTGTTTTTTTTTTTTAAATCAAA  
AAACAATTTTTTTTTTAAAAAAAGCTTTGAAAATGGATGGTTTGGGTATTAAAAAGAAAAAAACTTAAAAAA  
AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTTCCCTTGAGCAAAGCAGCCAGACGT  
TGAAGTGTGTTTTCTTTCCCTGGGCGCAGGGTGCCAGGGTGTCTTCCGATCTGGTGTGACCTTGGTCCAGGAGTT  
CTATTTGTTTCTGGGGAGGGAGGTTTTTTTGTGTTGTTTTTTGGGTTTTTTTTGGTGTCTTGTGTTTTCTTCTCCTC  
ATGTGTCTTGGCAGGCACTCATTTCTGTGGCTGTGCGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAG  
ACTCGGGTTGGCTAATCCCCGATGAACGGTGCTCCATTCGCACCTCCCCTCCTCGTGCCTGCCCTGCCTCTCCA  
CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTTTCCCACTCCTGCGGCATGGGTGTGT  
CCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCAGCCCTGTGCGCACCTGGTCTTTCATGAAGAGCAGACACTTA  
GAGGCTGGTCCGGAATGGGGAGGTGCGCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCTGGCGC  
CTGGAGTGACACAGCCAGTGGGCACTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT  
AGAAGGGTCCCCGCCTTAGATCAATCAGTGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGT  
CCATCCGTCTGTCCGTCCATTTGTGTTTTCTGCGTCGTGTGATGATATAATCCTCAGAAATAATGCACACTAG  
CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAAGTTGTAGACTCGGTACAGTATCAAATAAA  
ATCTATAACAGAAAAA

## **FIGURE 88**

MRQTIIKVIKFILIIICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI  
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK  
RFAVFLSEVSENKLRQLNLNNEWTLDKLRQLTKNAQDKLELHLFMLS GIPD TVFDLVELEV  
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLENLRLALHIKFTDIKEIPLWI  
YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI  
NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNHLKTIEEIIISFQ  
HLHRLTCLKLWYNHIAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYDLSSHNNLT  
FLPADIGLLQNLQNLAITANRIETLPPPELFQCRKLRLALHLGNNVLQSLPSRVGELTNLTQIE  
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

### **Transmembrane domain:**

amino acids 51-75 (type II)

### **N-glycosylation site.**

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 85-89

### **Casein kinase II phosphorylation site.**

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,  
398-402, 493-497

### **N-myristoylation site.**

amino acids 173-179, 261-267, 395-401, 441-447

## **FIGURE 89**

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGGCGCTCTCCCGT  
CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT  
GGCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG  
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCCTGGTCAT  
GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAACTTTGAGGAAATTGGGC  
CCCTTGACAGTGATCTCAAACCACGGAAAACCACTGGCTCCAGGCTGCCAGTCTCCTATTT  
GTGGATAATCCCGTGGGCACTGGGTTCA GTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA  
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG  
AATTCAGACAGTTCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC  
ATTGGTCTAGAGCTTTATAAGGCCATTCAAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT  
TGCCTTGGGTGATTCTCCTGGATCTCCCCTGTTGATTGGTGCTCTCCTGGGGACCTTACCTGT  
ACAGCATGTCTCTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA  
CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT  
GATCATTGAACAGAACACAGATGGGGTGAAGTCTATAACATCTTAACTAAAAGCACTCCCA  
CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC  
CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA  
GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA  
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC  
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTGAGGAGGCCTG  
GGTGCGGAACTGAAGTGGCCAGAACTGCCTAAATTAGTCAGCTGAAGTGAAGGCCCTGT  
ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC  
TACTGGATTCTGAAAGCTGGTCATATGGTTTCTTGACCAAGGGGACATGGCTCTGAAGAT  
GATGAGACTGGTGA CTGAGCAAGAAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT  
TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT  
AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT  
GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAACCTAAGATTTTAAAAAATTGAT  
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

## **FIGURE 90**

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC  
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSLKPRKTTWLQAASLLFVDNPVGTGFSY  
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR  
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVKNKGLYRE  
ATELWGKAEMIIIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQORDALS  
QLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL  
IVDTMGQEAWVRKCLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP  
SDQGDMALKMMRLVTQQE

### **Signal sequence:**

amino acids 1-25

### **N-glycosylation site.**

amino acids 64-68, 126-130, 362-366

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 101-105

### **Casein kinase II phosphorylation site.**

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

### **N-myristoylation site.**

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,  
187-193, 195-201, 331-337, 332-338, 360-366

## **FIGURE 91**

GGCCGCGGGAGAGGAGGCCATGGGCGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC  
GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGA  
CGGGTCATCACGTGCGGCATCGTGGGTGGAGAGGACGCCGAACCTCGGGCGTTGGCCGTGGCA  
GGGGAGCCTGCGCCTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG  
CACTCACGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG  
GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG  
TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG  
CCTTGGTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG  
GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA  
AGAGGATGAGGCACTGCCATCTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA  
ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG  
GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT  
GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG  
GTCGGCCCAATCGGCCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG  
CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT  
TCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTTGAGCCTACCTGAGCCCATGCAGCCTGGGGC  
CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTGGTAATAAACACATTCCAGTTGA  
TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

Sequence

## **FIGURE 92**

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW  
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI  
YLSPRYLGNSPYDIALVKLSAPVYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP  
SPHTLQEVQVAIINNSMCNHLFLKYFRKDI FGDMVCAGNAQGGKDACFGDSSGGPLACNKNG  
LWYQIGVVSWSGVGCRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL  
LGPV

### **Signal sequence:**

amino acids 1-18

### **N-glycosylation site.**

amino acids 167-171, 200-204, 273-277

### **Casein kinase II phosphorylation site.**

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

### **N-myristoylation site.**

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,  
259-265, 269-275

### **Amidation site.**

amino acids 33-37

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 252-263,

### **Serine proteases, trypsin family, histidine active site.**

amino acids 78-84

## FIGURE 93

Sequence

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCT  
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGC  
TGCCCCCAGGCTGGGTGTCCCTGGGCGGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT  
GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTGCGGATCCCAG  
CTCTCCTCAATACGGAATACTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC  
CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCAGAAAGTGCCATTCT  
GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC  
TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAACCCATGTTGTAAGGTCCCCAC  
ATCCCTACCAGCTTCCACAGGCCTTGGCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT  
TTTCCCCAACATCATCCCTGAGGCAACGTCTGAGCCGCAGGTGACAGGGACTGTAGGCCT  
GCATCTGGGGGTAACCCCTCTGTGATCCGTAAGCGATACAACTTGACCTACAAGACGTGG  
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCAGTTCTTGAGCAGTATTTCCATGAC  
TCAGACCTGGCTCAGTTCATGCGCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC  
CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCGGGATTGAGGCCAGTCTAGATGTGCAGT  
ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG  
GGACAGGAGCCCTTCCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT  
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA  
ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCTCAGGTGACAGT  
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG  
CCCCATGTCAACACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACAAATGAAA  
TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCACGGCCTTCATACCAGGAG  
GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCACCTGCCACCATCCAGTTACTTCAATGC  
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACCTTCTGATGGCTACTGGGTGGTCAGCAACA  
GAGTGCCCATTTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTGGGGGGATCCTA  
TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCTCTTGGCTTTCTCAACCCAAG  
GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC  
TGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCTGGCTGGGATCCTGTAACAGGC  
TGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCTATC  
AGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAGTTTCACTCCCTTATTCTGCCCTGTTG  
GAAGCCCTGCTGAACCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCCTGAAA  
TGCTGTGAGCTTGACTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT  
CCTGCCTTAGATTCTTCAATAAGATGCTGTAAGTAGCATTTTTTTGAATGCCTCTCCCTCCGC  
ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA  
TTTCACTTGATATTCAATCCCCAATTCAGTGCAAGGAGACCTCTACTGTCACCGTTTACTCT  
TTCCCTACCCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG  
GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAAGTTCTCTG  
ACTACTCTTGTCTTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC  
TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCTGGAACAAATCACTGACA  
TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA  
TGTAATAAA

## **FIGURE 94**

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS  
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL  
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGG LHRFPPTSSLRQRP  
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG  
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML  
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH  
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP  
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG  
RPPLGLFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

### **Signal sequence:**

amino acids 1-16

### **N-glycosylation site.**

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

### **Glycosaminoglycan attachment site.**

amino acids 361-365, 408-412, 538-542

### **Casein kinase II phosphorylation site.**

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

### **N-myristoylation site.**

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,  
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,  
521-527, 533-539, 549-555

**THE UNIVERSITY OF CHICAGO**

[illegible]

## **FIGURE 96**

MAGIPGLLFLLFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS  
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIIYLSSSGDGAQHRDSGSSGKS  
RRKRQIYGYSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG  
TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA  
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD  
AQPASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW  
IKGNYLDCREG

### **Signal sequence:**

amino acids 1-19

### **N-glycosylation site.**

amino acids 93-97, 207-211

### **Glycosaminoglycan attachment site.**

amino acids 109-113, 316-320

### **Casein kinase II phosphorylation site.**

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

### **N-myristoylation site.**

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

### **Serine proteases, trypsin family, histidine active site.**

amino acids 171-177

## **FIGURE 97**

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCGCCCCACCAGCCATGGTGGTTT  
CTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG  
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA  
GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA  
GCATCCAGAAGAATGGGACCCACCCTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC  
ACTGCTGCCCCACTGTTTCAAGGACAACCTGAACAAACCATACCTGTTCTCTGTGCTGCTGGG  
GGCCTGGCAGCTGGGGAACCCCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC  
CCCACCCTGTGTATTCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCCTCTCGAG  
CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT  
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC  
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTTCTATCATCGACTCGGAAGTCTGCAGCCAT  
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT  
GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG  
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC  
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA  
GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG  
CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA  
TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACC  
TCTACCTCTGGGGGCCCCGACGGCTGCTGCGGAAAGGAAACCCCTCCCCGACCCGCCCCGAC  
GGCCTCAGGCCCCCCTCCAAGGCATCAGGCCCCGCCAACGGCCTCATGTCCCCGCCCCCAC  
GACTTCCGGCCCCCGCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT  
AGGTATTTGTAACCCTGCCACATATCTTATTTATTCCTCCAATTTCAATAAATTATTTATT  
CTCCAAAAAAAAA

## **FIGURE 98**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP  
WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA  
WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG  
VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC  
QVDGAWLLAGIISWGECAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG  
SGAAARS

### **Signal sequence:**

amino acids 1-32

### **N-glycosylation site.**

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,  
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

### **Glycosaminoglycan attachment site.**

amino acids 826-830

### **Casein kinase II phosphorylation site.**

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,  
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,  
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,  
1013-1017, 1019-1023, 1021-1025

### **Tyrosine kinase phosphorylation site.**

amino acids 607-615

### **N-myristoylation site.**

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,  
612-618, 623-629, 714-720, 873-879

[illegible]

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCTCTGCTGCCGCTACTGCTA  
CTGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGTATGGT  
GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA  
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCCTGTGGGGCCAC  
AACAAAGGAGCGCGGGCGCCGCGGCGAGAATCTGTTGCCATCACAGACGAGGGCATGGACGT  
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT  
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC  
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT  
GGTGTGCAACTATGAGCCTCCGGGGAAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC  
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC  
CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA  
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT  
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAAACCCAGGCC  
CCAACCTTCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC  
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGATGAGGAGCCAGTTA  
CCTTCCCCAAATCGACCCATGTTCCCTATCCCAAATCAGCAGACAAAGTGACAGACAAAACA  
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG  
GGAACCTCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG  
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC  
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACTCTGCCACCGCTAA  
TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCTTGCCAGGTGCAGAGGGCCCTGACA  
AGCCTAGCGFTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA  
CTACTGCTCCTGCCTCCTCTGGTGTGGCTGGAATCTTCTTGAATGGGATAACCACTCAAAGGG  
TGAAGAGGTCAGCTGTCCTCCTGTCTATCTTCCCCACCCTGTCCCCAGCCCCCTAAACAAGATA  
CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC  
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACC  
GGGCCCACACCTCTCCTGCCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT  
CACTGCCTACCTGGCCTGGGGCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG  
TGTAGCTGGGGATGGGGATTCCCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC  
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT  
GTCCAACCTGTGAAA

## **FIGURE 100**

MHGSCSFLMLLLPLLLLLVATTGPVGALTDEEKRLMVELHNLVYRAQVSPTASDMLHMRWDEE  
LAFAKAYARQCVWGHNKERGRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ  
MCGHYTQVWAKTERIGCGSHFCEKLQGVETNIELLVVCNYEPPGNVKGKRPYQEGTPCSQC  
PSGYHCKNSLCEPIGSPEDAQDLPLYLVTEAPSFRAEASDSRKMGTTPSSLATGIPAFLVTEV  
SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS  
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEELPPSSEVLAS  
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV  
SGLNSGPGHVWGPLLGLLLLPLVLGIF

**Signal sequence:**

amino acids 1-22

**N-glycosylation site.**

amino acids 114-118, 403-407, 409-413

**Glycosaminoglycan attachment site.**

amino acids 439-443

**Casein kinase II phosphorylation site.**

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

**N-myristoylation site.**

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,  
250-256

**Amidation site.**

amino acids 82-86, 172-176

**Peroxidases proximal heme-ligand signature.**

amino acids 287-298

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.**

amino acids 127-138

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.**

amino acids 160-172

[illegible][illegible]

## FIGURE 102

MVDVLLLFSLCLLFHISRPLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPV SAN  
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP  
GYFDNLANTLLVLKLNRRNRI SAIPPKMFKL PQLQHLELNRNKIKNVDGLTFQGLGALKSLKM  
QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE  
FCQKLSELDLTFNHL SRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE  
ISWTIEDMNGAFSGLDKLRRLLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ  
MKKLQQLHLNTSSLLCDCQLKWL PQWVAENN FQSFVNASCAHPQLLKGRSIFAVSPDGFVCD  
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENY AHLRAQG  
GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA  
MARLECAAVGHPAPQIAWQKDG GTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN  
SAGSISANATLTVLETPSFLRPLLDRTVT KGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER  
HFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDG  
WATVGVVIIIAVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDET NLPADIPSYLSSQGT LAD  
RQDGYVSSES GSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP  
MYLKGNVYGSDPFETYHTGCSPDPRTV LMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW  
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDA  
YSSFGQPSDCQPRAFYLKAHSSPDLD SGSEEDGKERTDFQEENHICTFKQTL ENYRTPNFQS  
YDLDT

### Signal sequence:

amino acids 1-19

### Transmembrane domain:

amino acids 746-765

### N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,  
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

### Glycosaminoglycan attachment site.

amino acids 826-830

### Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,  
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,  
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,  
1013-1017, 1019-1023, 1021-1025

### Tyrosine kinase phosphorylation site.

amino acids 607-615

### N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,  
612-618, 623-629, 714-720, 873-879

**THE** **NEW** **YORK** **PUBLIC** **LIBRARY**

GGGGAGAGGAATTGACCATGTAAAGGAGACACTTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAAATG  
 AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGAAC  
 GAAGCTTTTTCTTGTGAGCCCTGGATCTTAAACACAAATGTGTATATGTGCACACAGGGAGCATTCAGAATGAAA  
 TAAACACAGAGTTAGACCCGCGGGGGTTGGTGTGTTCTGCATAAAATAAATAATCTTAAAGCAGCTGTTCCCTTCC  
 CCACCCCCAAAAAAGGATGATTGAAATGAAGAACCAGGAGATTCAAAAGAAAAAAGTATGTTTCATTTTTCTC  
 TATAAAGGAGAAAGTGACCAAGGAGATATTTTTTGGAAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT  
 GGTGTGGTGGTGTTCCTTTCTTTTTGAATTTCCCAAGAGGAGAGGAATTAATAATACATCTGCAAGAA  
 TTTTTCAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCACAGTTGGA  
 TTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTTTTTTTAAAT  
 TTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTCTTAAACCACCTGGATTTCATCTGGATGTTGCT  
 GTGATCAGTCTGAAATCACACTGTTTGAATTCAGAAGGACCAACACAGATAAATTATGAATGTGCAACAGAT  
 GACCTTACATCCACAGCAGATAATGTATAGGTCCTAGGTTTAAACAGGGCCCTATTTGACCCCTGCTTGTGGTGCT  
 GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA  
 CCAGTTTCAGCAAGGTGATTGTGTTTCGAAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT  
 GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT  
 ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATGGGGCTTTCAATGGCTGGCGAACCTCAACACTCTGGA  
 ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATCTTGCTTAAACTGAAGGAGCTGTGGT  
 GCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAAACAGAATTCCTTCTTTGCGCCGACTAGACTTAGG  
 GGAATTGAAAAGACTTTTATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACCTTGAGGTATTTGAACCTTGC  
 CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAACTAGATGAGCTGGATCTTTCTGGGAATCA  
 TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGACCTTCAAAAACCTGTGGATGATACAGTCCCAGAT  
 TCAAGTGATTGAACCGAATGCCTTTGACAACCTTTCAGTCACTAGTGAGATCAACCTGGCACAACAATAATCTAAC  
 ATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCAACCCCTTGAACATG  
 TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTTGTGCCCGGTG  
 TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCATGCTATGCTCCGGT  
 GATTGTGGAGGCCCTTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTGCGGCCTCCACATG  
 CTGCATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACATCATGGGCGGTACAAGTGCGGATAGCTGT  
 GCTCAGTGATGGTACGTTAAATTTCACAAAATGTAACGTGTCAGAGATACAGGCATGTACATGTATGGTGAGTAA  
 TTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCAGCAACCACTACTCCTTTCTTACTTTTC  
 AACCCTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCC  
 AGTGGTGCAGTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTGCAGAGAGAAAACCTT  
 CACCATCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTCATGAAGACTACCAAATCATCAT  
 TGGGTGTTTTTGGCCATCACACTCAGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA  
 TCGGCAAAACCATCACGCCCCAACAAAGGACTGTTGAAAT'TATTAATGTGGATGATGAGATTACGGGAGACACACC  
 CATGGAAGCCACCTGCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT  
 CAACCACACAACACAGTTAAACACAATAAATTCAATACACAGTTTCAGTGCATGAACCGTTATTGATCCGAATGAA  
 CTCTAAAGACAATGTACAAGAGACTCAAATCTTAAACATTTACAGAGTTACAAAAAACAAACAATCAAAAAA  
 CACAGTTTTATTAAAAATGACACAATGACTGGGTCAAATCTACTGTTTCAAAAAGTGTCTTTACAAAAAACA  
 AAAAGAAAAGAAATTTATTTATTAAAAATCTATTGTGATCTAAAGCAGACAAAAA

## **FIGURE 104**

MLNKMTLHPQQIMIGPRFNRAFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC  
VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA  
NLNTLELFDNRLTTIPNGAFVYLSKLELWLRNPNIESIPSYAFNRIPSLRRLDLGELKRLS  
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLELDLSGNHLSAIRPGSFQGLMHLQKL  
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNCDIL  
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE  
LKCRASSTLSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTQDTGMYTCMVNSVGN  
TTASATLNVTAATTTTFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSSTLPQ  
STRSTEKFTTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQN  
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS  
VHEPLLIRMNSKDNVQETQI

### **Signal sequence:**

amino acids 1-44

### **Transmembrane domain:**

amino acids 523-543

### **N-glycosylation site.**

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,  
442-446, 488-492, 606-610

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

### **Casein kinase II phosphorylation site.**

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

### **N-myristoylation site.**

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,  
391-397, 422-428, 433-439, 531-537

# FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGCA  
TGAAAGGGCTTCGCCGCCGGGAGTAAAAGAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC  
GAGGGCGGGCGTGACCCCTCGGCTGGAAGTTTGTGCCGGGCCCCGAGCGCGCGCGGGCTGGGAGCTTCGGGTAGA  
GACCTAGGCCGCTGGACCGCGATGAGCGCGCCGAGCCTCCGTGCGCGCGCGCGGGGTGGGGCTGCTGCTGTGC  
GCGGTGCTGGGGCGCGCTGGCCGGTCCGACAGCGGCGGTGCGCGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC  
GAGCGCCCATGCCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGCTTT  
CCCGAGCCACTCCCGTCCCTGGGTGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTTCATCAAGGCAAGTTCC  
ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA  
GTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG  
TTTCAGTCCCTTGAACCTTTGGACCTTAGCAGCAACAATATTTAGAGCTCCAACTGCATTTCCAGCCCTACAG  
CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA  
CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAAGACTTTGGATCTGAAGA  
CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT  
CTGAAAATGCAAAGAAATGGAGTAACGAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTG  
CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACCTCAT  
CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA  
ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG  
AACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC  
AATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAACTGAGGCGACTGATACTC  
CAAGGAAATCGGATCCGTCTTATTAATAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT  
GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCAAAATGAAGAAACTGCAACAATTGCATTTAAATACA  
TCAAGCCTTTTGTGCGATTGCCAGCTAAATGGCTCCACAGTGGGTGGCGGAAAACAACCTTTCAGAGCTTTGTA  
AATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTGTCTGTTAGCCAGATGGCTTTGTGTGT  
GATGATTTTCCAAACCCAGATCACGGTTAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC  
ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGTCTGGAAAAAGACAATGAACTACTGCATGAT  
GCTGAAATGGAAATATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG  
CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCATCTCCAATCACTTTGGTTTCATCTACTCTGTC  
AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC  
ATGGCAGCTTGGAGTGTGCTGTGTTGGGGCACCAGCCCCCAGATAGCCTGGCAGAAAGGATGGGGGCAAGAC  
TTCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCCAGGATGACGTGTTCTTTATCGTGGATGGAAGATA  
GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTAGCAAAATGCAACTCTGACTGTC  
CTAGAAACACCATCATTTTTTGGGCCACTGTTGGACCGAAGTGAACCAAGGGAGAAACAGCCGTCTACAGTGC  
ATTGCTGGAGGAAGCCCTCCCCCTAACTGAACCTGGACCAAAGATGATAGCCCATTTGGTGGTAACCGAGAGGCAC  
TTTTTTGCTAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTGATGCTGGGAAATACACATGTGAG  
ATGTCTAACACCCCTTGGGAGTGTGCTGTGTTGGGGCACCAGCCCCCAGATGATGCTGATCCCACTCCAACCTGCGACTCCCCT  
CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGTGT  
GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACC  
AACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCTCATCTCAGGGAACGTTAGCTGACAGGCAG  
GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACAGTTTGTGCATCTTCAGGTGCTGGATTTTTCTTACCA  
CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTT  
CTTTGTCCGTTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGAAATGTGTATGGCTCAGATCCTTTTGAACA  
TATCATACAGGTTGCAGTCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAG  
GAGTGCTACCCATGTTCTCATCCTTCAGAAAGATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT  
GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC  
TCTTTAGATTTTAGTGCAATCCAGAGCCAGCGTGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA  
GCTCTCAGGAGACCTTACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTGAGCCAAGAGCCTTTTAT  
TTGAAAGCTCATTCTTCCCAGACTTGGACTCTGGGTGAGAGGAAGATGGGAAAGAAAGGACAGATTTTTCAGGAA  
GAAATCAGATTTGTACCTTTAAACAGACTTTAGAAAACCTACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC  
ACATAGACTGAATGAGACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACCTTTTATTTAAAGAGAGAGAAT  
CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAAATGCTTTATTTATACAGATGAACCAAATTAAC  
AAAAAGTTATGAAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAACTATTTTTTAACTTTG  
TTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTTTATAATGCCAGA  
TTTCTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAAATAGAAGTT  
ACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1000 900 800 700 600 500 400 300 200 100 0

[illegible]

**Signal sequence:**

**Transmembrane domain:**

N-glycosylation site.

Glycosaminoglycan attachment site.

Casein kinase II phosphorylation site.

**Tyrosine kinase phosphorylation site.**

N-myristoylation site.

Leucine zipper pattern.

amino acids 58-80, 65-87

STATION	DATE	TIME	WIND	TEMP	REL	WIND	TEMP	REL
1	10/10/10	0000	000	10.0	100	000	10.0	100
2	10/10/10	0100	000	10.0	100	000	10.0	100
3	10/10/10	0200	000	10.0	100	000	10.0	100
4	10/10/10	0300	000	10.0	100	000	10.0	100
5	10/10/10	0400	000	10.0	100	000	10.0	100
6	10/10/10	0500	000	10.0	100	000	10.0	100
7	10/10/10	0600	000	10.0	100	000	10.0	100
8	10/10/10	0700	000	10.0	100	000	10.0	100
9	10/10/10	0800	000	10.0	100	000	10.0	100
10	10/10/10	0900	000	10.0	100	000	10.0	100
11	10/10/10	1000	000	10.0	100	000	10.0	100
12	10/10/10	1100	000	10.0	100	000	10.0	100
13	10/10/10	1200	000	10.0	100	000	10.0	100
14	10/10/10	1300	000	10.0	100	000	10.0	100
15	10/10/10	1400	000	10.0	100	000	10.0	100
16	10/10/10	1500	000	10.0	100	000	10.0	100
17	10/10/10	1600	000	10.0	100	000	10.0	100
18	10/10/10	1700	000	10.0	100	000	10.0	100
19	10/10/10	1800	000	10.0	100	000	10.0	100
20	10/10/10	1900	000	10.0	100	000	10.0	100
21	10/10/10	2000	000	10.0	100	000	10.0	100
22	10/10/10	2100	000	10.0	100	000	10.0	100
23	10/10/10	2200	000	10.0	100	000	10.0	100
24	10/10/10	2300	000	10.0	100	000	10.0	100

CAAAACCTTGCCTCGCGGAGAGCGCCACGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC  
TGGGGGAGCGCGTTTCGGCTGTGGGGCGCCGCTCGGCGCCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC  
CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCCTGGGTGGTCCCCTCCCCTATCCCTCCTTTATATA  
GAAACCTTCCACACTGGGAAGGCAGCGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG  
GCGCACAGCAATCCGAGTTTACAGATTTTTACAGATACCAAATGGAAGGCGAGGAGGCAGAACAGCCTGCGCTGGT  
TCCATCAGCCCTGGCGCCCGCAGCTGACCTCGGCACCTCTGCAGGCACCATGGCCACAGACCGGGTGCTGCTGC  
TGCTCCTGCTGCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGGCTGAGGGCCCCAGGATTGGCCGAA  
GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGCTGGTACTGAGCCCTGAGG  
AGCCCGGGCCTGGCCAGCCGCGGTGAGTGCCCCCGAGACTGTGCCTGTTCCAGGAGGGCGCTCGTGGACTGTG  
GCGGTATTGACCTGCGTGAGTTCCCGGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACCAACCAGC  
TGAAAAAGATCTACCTTGAGGAGCTCTCCCGGCTGCACCGGTGGAGACACTGAACCTGCAAAACAACCGCCTGA  
CTTCCCGAGGGCTCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTTGGCCAATAACAAGC  
TGACCTTGGCACCCCGCTTCTGCCAAACGCCCTGATCAGTGTGACTTTTGTGTCGCAACTATCTACCAAGATCT  
ATGGGCTCACCTTTGGCCAGAAGCCAAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCGGGC  
TGCCGGACAACATGTTCAACGGCTCCAGCAACGTCGAGGTCTCATCTGTCCAGCAACTTCTGCGCCACGTGC  
CCAAGCACCTGCCGCTGCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT  
TCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGGCCTGGACAACGAGA  
CCTTCTGGAAGCTCTCAGCCTGGAGTACCTGGATCTGTCCAGCAACACCTGTCTCGGGTCCCAGCTGGGCTGC  
CGCGCAGCCTGGTGCTGCTGCACTTGGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCCATCC  
GCAGCCTGGAGTACCTGCTGCTGCACAGCAACAGCTGCGGGAGCAGGGCATCCACCACTGGCCTTCCAGGGCC  
TCAAGCGGTTGCACACGGTGCACCTGTACAACAACGCGCTGGAGCGCGTGCCAGTGGCCTGCCTGCGCCGCTGC  
GCACCCTCATGATCTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCTTGAGG  
AGCTCAACCTCAGCTACAACCGCATCACACGCCACAGGTGCACCCGCGACGCCTTCCGCAAGCTGCGCCTGTGCTGC  
GCTCGCTGGACCTGTGCGGCAACCGGTGCACAGCTGCCACCTGGGCTGCCCTCGAAATGTCCATGTGCTGAAGG  
TCAAGCGCAATGAGCTGGCTGCTTTGGCAGAGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACTTCA  
CCAGCAACCGACTGCGCAGCCGAGCCCTGGGGCCCCGTGCCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA  
TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAA  
AGATTAGTGCGGTGCCCCCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAAACAAGC  
TGGGTGTGGGCTCCGTGGTGGAAGCTGCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGAACATTGAAGGCAACT  
TAGAGTTTGGTGACATTTCCAAAGGACCGTGGCCGCTTGGGGGAAGGAAAGGAGGAGGAGGAAGGAGGAGGAGG  
AGGAAGAGGAAACAAGATAGTGAACAAGGTATGTCAGATGTACCTAGGATGATGGACGCGGACTCTTTCTGC  
AGCACACGCCTGTGTGCTGTGAGCCCCCATCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCA  
TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCACGGCGTGTCCACGGCCAGACACATGC  
ACACACATCACACCCTCAAACACCCAGCTCAGCCACACACAACCTACCCTCCAAACCACCACAGTCTCTGTACAC  
CCCCACTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGGTCTGCCCTGCCCTGGCACACAGGCACCCA  
TTCCTCCCCCTGTGACATGTGTATGCGTATGCATACACACCAACACACACATGCACAAGTCAATGTGCGAA  
CAGCCCTCCAAAGCCTATGCCACAGACAGCTCTTGCCCCAGCCAGAATCAGCCATAGCAGCTGCGCGTCTGCCCT  
GTCCATCTGTCCGTCCGTTCCTTGGAAGACACAAGGGTATCCATGCTCTGTGGCCAGGTGCCTGCCACCTCT  
GGAACTCACAAAAGCTGGCTTTTATTCTTTCCATCCTATGGGGACAGGAGCCTTCAGGACTGCTGGCCTGGCC  
TGGCCACCCTGCTCCTCCAGGTGCTGGGCAGTCACTCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA  
CAGGCACCTTTTCCAATGGGCAAGCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGCTGGGGCTTTGGG  
CAGGAGTGAAGCAGAGGTGATGGGCTGGGCTGAGCCAGGGAGGAAGGACCAGCTGCACCTAGGAGACACCTTT  
GTTCTTCAGGCTGTGGGGGAAGTTCCGGGTGCTTTATTTTATTTCTTTCTAAGGAAAAAATGATAAAAAAT  
CTCAAAGCTGATTTTTTCTTGTATAGAAAACTAATATAAAAGCATTATCCCTATCCCTGCAAAAAA

## **FIGURE 108**

MEGEEAEQPAWFHQWRPGASDSAPPAGTMAQSRVLLLLLLLPPQLHLGPVLAVRAPGFGRS  
GGHSLSPREENEFAEEEPVLVLSPPEPGPGPAAVSCPRDCACSQEGVVDCCGIDLREFPGDLP  
EHTNHLSQLNNQLEKIYPEELSRHRLETLNLQNNRLTSRGLPEKAFFEHLTNLNYLYLANNK  
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRVYLHNNKLADAGLPDNMFNGSSNV  
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN  
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSN  
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMLHNQITGIGREDFATTYF  
LEELNLSYNRITSPQVHRDAFRKLRLRLSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA  
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL  
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD  
RGRLGKEKEEEEEEEEEEEEEETR

### **Signal sequence:**

amino acids 1-48

### **N-glycosylation site.**

amino acids 243-247, 310-314, 328-332, 439-443

### **Casein kinase II phosphorylation site.**

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

### **N-myristoylation site.**

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,  
477-483, 498-502, 539-545, 548-554

### **Leucine zipper pattern.**

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,  
535-557

## FIGURE 109

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTCCTCCGGGAGCGGCAG  
CAGTAGCCCCGGGCGGCGAGGGCTGGGGGTTCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCCACCACC  
CAACCTGTTCTCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCAACATGGATTTTCTCCTGGCGCTGGT  
GCTGGTATCCTCGCTCTACCTGCAGGCGGCGCCGAGTTCGACGGGAGGTGGCCAGGCAAATAGTGTATCGAT  
TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTAGCC  
TGTGTGCCAACACCATGCAAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCCTGGTTATGCTGG  
AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA  
CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGATGGTTCCTGCTCAAGTGCCCTGACCTG  
CTCCATGGCAAACCTGTAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCCTGGCCT  
GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT  
TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCTATAAGGCTTCGATCTCATGTATATTGGAGG  
CAAATATCAATGTGATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA  
CGTACGTGGGTCTTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTGTGTGTATATCCAAA  
AGTTATGATTGAACCTTCAGGTCCAATTGATGTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA  
TAATAATTGGATTCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA  
CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCTACTCCACCACCACCACCACC  
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACACCACCGGACTGACAACTATAGC  
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCTCAGAAACCCAGAGG  
AGATGTGTTTCACTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAA  
TGACTTGCAGTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCCGCAGCCAAAGCCCCAGG  
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCTCATGCATTGAGGGGACCTGTGCCTGTCATTGAGGCA  
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAAACCGGTGCCACGGAGCAGCCCTGTG  
GGGAAGAAATGGTGGCCATGGCTGGAGGCAAACACAGATCACTTGCAGGGGGCTGACATCAAGAGCGAATCACA  
AAGATGATTAAAGGGTTGGAAAAAAGATCTATGATGGAATAAAGGAACCTGGGATTATTGAGCCTGGAGAAG  
AGAAGACTGAGGGGCAAACCATGATGGTTCAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG  
TTCTCCATATGCACTAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGG  
GGCCATTGTTAGAATACTTCATAAAAAAAGAAAGTGTGAAATCTCAGTATCTCTCTCTTTCTAAAAAATTAGA  
TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCAAA  
AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG  
CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCACCAAATGCTGAGCTCACTGAAATA  
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAACTATTTATTCCAAATGAGAGTATGATGGAC  
AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC  
TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTTAAGAGATCCTTCAAGGAACACAGTTCAGAGAG  
ATTTTCATCGGGTGCACTCTCTCTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC  
ACACCGGCAGACCTTTTCCCTCACCTCATCAGTATGATTGATTTCTCTTATCAATGGACTCTCCAGGTTCCAC  
AGAACAGTAATATTTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCAATTAACTGGTAAAGGCAGGGCTGG  
AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTAAATGGTTTATT  
TCCTTTATGGTCATATAACTGCACAGCTGAAGATGAAAGGGGAAAATAAATGAAAATTTTACTTTTCGATGCCAA  
TGATACATTGCACTAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTATTATTATTTAATGTTTT  
CTAAAAATAAAAAATGTTAGTGGTTTTTCCAAATGGCCTAATAAAAACAATTATTTGTAAATAAAAAACTGTTAGTAAT

1000 900 800 700 600 500 400 300 200 100 0

## **FIGURE 110**

MDFLLALVLVSSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP  
RCKHGEICIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD  
GSCSSALTCSMANCQYGCDVVKGQIRCQCPSPLHLAPDGRTCDVDVDECATGRASCPRFRQC  
VNTFGSYICKCHKGFDLMYIGGKYQCHDIDEC SLGQYQCSSFARCYNVRGSYKCKCKEGYQG  
DGLTCVYIPKVMIEPSGPIHVPKGN GTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRP  
TSKPTTRPTPKPTPIPTPPPPPLPTLRTPLPPTTPTTGLTTIAPAASTPPGGITVDN  
RVQTD PQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGG  
KAARLVLP LGRMLHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI  
TLRGADIKSESQR

### **Signal sequence:**

amino acids 1-17

### **N-glycosylation site.**

amino acids 273-277

### **Casein kinase II phosphorylation site.**

amino acids 166-170, 345-349

### **Tyrosine kinase phosphorylation site.**

amino acids 199-206

### **N-myristoylation site.**

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,  
421-427, 433-439, 462-468, 476-482

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 104-116, 186-198, 231-243

### **Cell attachment sequence.**

amino acids 382-385

### **EGF-like domain cysteine pattern signature.**

amino acids 75-87

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCTTTAGATTGTGA  
AATGTGGCTCAAGGTCTTCACAACCTTTCTTTCTTTGCAACAGGTGCTTGCTCGGGGCTGA  
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC  
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA  
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTTCTGACTTGGAATACC  
AACACAAGTTTACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT  
GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACCTATCTGCCAGTCA  
GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT  
CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG  
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC  
TCCCCAAAACAATAACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT  
GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAGTGATATCATTATGCCCATCATATATTAT  
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT  
TGACCTTGAGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT  
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT  
GCATCTGAGAAAGTAGCCCAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC  
CGGCAGGCAAGATGAAACTCATTTACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG  
CACAGAAAGGAAAATCATTGTCACTTTTAGCAAGTATAACTGGAATATCACTATTTTTTGATT  
ATATCCATGTGTCTTCTCTTCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAA  
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG  
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG  
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT  
GTATGAAGTTATTTCAGCACATCCCTGCCCAGCAGCAAGACCATCCAGAGTGAACCTTTCATGG  
GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT  
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA  
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT  
GTGCAACAGAAAAACATGTTGGGGAAATATTCTCAGTGGAGAGTCGTTCTCATGCTGACGG  
GGAGAACGAAAGTGACAGGGGTTTTCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA  
ATTAGTTCTACTCTACACTTTCATCTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA  
TGTGGAAACTTTTACATTGTTTCGATTTTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG  
TTAAGAATGCTAAATTTATGTTTCAATTTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA  
CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACCTATGCCTTCTCTTTTTTTTCAATCACC  
AGTAGTATTTTTTGAGAAGACTTGGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTTA  
TTTTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTTTGCTTTTAAAAA

## **FIGURE 112**

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH  
TMPKYLLGSVNKSVVPDLEYQHKFTMMPNALLINPLQFPDEGNYIVKVNIQNGTLSASQ  
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNRGPVHTSSTYSFS  
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV  
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT  
GRQDETHFTVITTSVGLEKLAQKGKSLSPASITGISLFLIISMCLLFLWKKYQPYKVIKQK  
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRVSPASDCVSGQDLHSTV  
YEVIQHIPAQQDHPHPE

### **Signal sequence:**

amino acids 1-18

### **Transmembrane domain:**

amino acids 341-359

### **N-glycosylation site.**

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,  
276-280, 308-312

### **Casein kinase II phosphorylation site.**

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

### **Tyrosine kinase phosphorylation site.**

amino acids 272-280

### **N-myristoylation site.**

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,  
239-245

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

112

**THE UNIVERSITY OF CHICAGO**

**THE UNIVERSITY OF CHICAGO**

[illegible]

amino acids 1-26

## amino acids 182-201

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

## amino acids 107-115

amino acids 20-26, 192-198

amino acids 25-29

## FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTTCGTAACTAATTCAACAAACGGGACCCTT  
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCAGTGGGACAGGCGGATTGGAAGAGCGGG  
AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAACTCTGGGTGTCTGC  
ATTGCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCCTCACCTCTATTG  
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC  
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC  
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAACTGG  
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGCAGGACTCAGCTGCA  
GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG  
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA  
GAGGGGAACCTTCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG  
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATACGGTGTGTGTGGATGGAGCAGGTGCT  
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA  
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCTGCTC  
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGGAATCTGCGGTACTTTGAGCAGTTATTGGA  
GGAAGAGAGAGAAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCCAAGGCA  
TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGTGGG  
GAGGGTGTCAAACCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA  
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA  
TCGTCAGGTACTACGATGTCTGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA  
CCTAAACTTGCACGAGCCACCGTTCTGTGATCCCAAGACAGGAGTCCTCACTGTCTGCCAGCTA  
CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC  
GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT  
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG  
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG  
GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG  
TTCTGGTACAACCTCTTGCGGAGCGGGGAAGGTGACTACCGAACAAGACATGCTGCCTGCCC  
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT  
TGAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTTCTGTCTTCCCCTTCTGCTC  
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCTTTGTATGTTCTTATCAGGCT  
GATTTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT  
GTGACTGAAGTCCCAGCCCTTCCATTAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA  
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT  
TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTGTTTTTAT  
GGCATTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAA

## FIGURE 116

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA  
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFP  
TDEDEIGAALMRLQDTYRLDPGTISRGLPGTKYQAMLSVDDCFGMGRSAYNEGDDYHTV  
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSHERAGGNLR  
YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTTPRRQKRLF  
CRYHHGNRAPQLLIAPFKEDEWDSPHIVRYDVMSDEEIERIKEIAKPKLARATVRDPKTG  
VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS  
RRPFDGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGECDYR  
TRHAACPVLVGCKWVSNKWFHERGQEFLRPGSTEVD

**Signal sequence:**

amino acids 1-17

**N-glycosylation site.**

amino acids 115-119, 264-268

**Glycosaminoglycan attachment site.**

amino acids 490-494

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 477-481

**Casein kinase II phosphorylation site.**

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,  
346-350, 365-369, 385-389, 457-461, 530-534

**Tyrosine kinase phosphorylation site.**

amino acids 71-80, 489-496

**N-myristoylation site.**

amino acids 14-20, 131-137, 171-177, 446-452

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 8-19

**Leucine zipper pattern.**

amino acids 213-235

## FIGURE 117

GCAGTATTGAGTTTTACTTCCTCCTCTTTTTAGTGGAAGACAGACCATAATCCCAGTGTGAGTGAAATTGATTGT  
TTCATTTATTACCGTTTGGCTGGGGGTAGTTCCGACACCTTCACAGTTGAAGAGCAGGCAGAAGGAGTTGTGA  
AGACAGGACAATCTTCTTGGGGATGCTGGTCCCTGGAAGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCC  
CAGGTTCTCTGGTTAAACTGAAAGCCTACTACTGGCCTGGTGCCCATCAATCCATTGATCCTTGAGGCTGTGCC  
CCTGGGGCACCCACCTGGCAGGGCCTACCACCATGCGACTGAGCTCCCTGTTGGCTCTGCTGCGGCCAGCGCTTC  
CCCTCATCTTAGGGCTGTCTCTGGGGTGACGCTGAGCCTCCTGCGGGTTTCTGGATCCAGGGGAGGGAGAAG  
ATCCCTGTGTGAGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTGAGAGCTCGGCTAGACCAAAGTG  
ATGAAGACTTCAAACCCCGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC  
GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTGTCTGGTGGCTGTCTGACCTCCCGAGCTACACTGTCCA  
CTTTGGCCGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTACTCTACTTCACTGGGCAGCGGGGG  
CCCGGGCTCCAGCAGGGATGAGGTGGTGTCTCATGGGGATGAGCGGCCCCGCTGGCTCATGTGAGAGACCTGC  
GCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGCAGGCCC  
CCCGCTGGCAGCCCTTGCTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTCAATTG  
GCGCAGGCGAGCAGGCCCGGTACTGTCTATGGGGCTTTGGCTACCTGTTGTACGGAGTCTCCTGCTTCGTCTGC  
GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCTGACGAGTGGCTTGGACGCTGCCTCATTG  
ACTCTCTGGGCGTCGGCTGTGTCTCAGACACCAGGGGCAGCAGTATCGCTCATTTGAACTGGCCAAAAATAGGG  
ACCCTGAGAAGGAAGGGAGCTCGGCTTTCCTGAGTGCCTTCGCCGTGCACCCTGTCTCCGAAGGTACCCTCATGT  
ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGA  
TCCGGAACCTGACCGTGCTGACCCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCGTTGGGCTCCCTGCTCCTTTCA  
CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCCTGTGCAGATGGGG  
CTCCCAAGTGCCACTACAGGGGGCTAGCAGGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA  
ATCGGCGCTATCAGCCCCGCTGCGCTTCCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC  
GGGGCATGGAGTACACCTGGACCTGCTGTTGGAATGTGTGACACAGCGTGGGCACCGCGGGGCCCTGGCTCGCA  
GGGTGAGCTGCTGCGGCCACTGAGCCGGGTGGAATCCTACCTATGCCCTATGTCACTGAGGCCACCCGAGTGC  
AGCTGGTGCTGCCACTCCTGGTGGCTGAAGCTGTGACGCCCGGCTTTCCTCGAGGCGTTTGACGCCAATGTCC  
TGGAGCCACGAGAACATGCATTGCTCACCCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG  
ACCCATTTCTTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTGGCTCG  
CTGTGCGAGCAGAGGCCCTTCCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCTGTGGACACTCTCT  
TCTTCCTTACCACCGTGTGGACAAGGCCTGGGCCCCGAAGTCTCAACCGCTGTGCGATGAATGCCATCTCTGGCT  
GGCAGGCCTTCTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTACCAACAGAGATCACCCCCAGGGCCCC  
CGGGGGCTGGCCCTGACCCCCCTCCCCCTCCTGGTGCTGACCCCTCCCGGGGGGCTCCTATAGGGGGGAGATTTG  
ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACGCTGACTACCTGGCGGGCCGAGCCCCGGCTGGCAGGTGAAC  
TGGCAGGCCAGGAAGAGGAGGAAGCCCTGGAGGGGCTGGAGGTGATGGATGTTTTCTCCGGTTCTCAGGGCTCC  
ACCTCTTTCCGGGCCGTAGAGCCAGGGCTGGTGACAGAAGTTCTCCCTGCGAGACTGCAGCCCACGGCTCAGTGAAG  
AACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGCTAGGGGGCCGTGCCAGCTGGCTATGGCTCTCTTTG  
AGCAGGAGCAGGCCAATAGCACTTAGCCCCGCTGGGGGCCCTAACCTCATTACCTTCTCTTGTCTGCCTCAGCC  
CCAGGAAGGGCAAGGCAAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTTTAAATATGAAAATGTTATTAA  
ACATGTCTTCTGCC

## FIGURE 118

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDNRRLD  
QSDDEFKPRIVPYRDPNKPYPKKVLRTRYIQTELGSRRERLLVAVLTSRATLSTLAVAVNRTV  
AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFFIMQDDTY  
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRLLLLRLRPHLDGCRG  
DILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE  
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV  
LGWDYFTEQHTFSCADGAPKCPLOQASRADVGDALETALEQLNRRYQPRLRQKQRLNNGYR  
RFDPARMEYTLDDLLECVTQRGHRRALARRVSLRPLSRVEILPMPYVTEATRVQLVLPPL  
VAEAAAAPAFLEAFAANVLEPREHALLTLLLTVYGPREGGRGAPDPFLGVKAAAELERRYPG  
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFPP  
VHFQEFNPALSPQRSPPGPPGAGDPDPSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLAA  
RARLAGELAGQEEEEALEGLEVMDFLRFSGLHLFRAVEPGLVQKFSRLRDCSPRLSEELYHR  
CRLSNLEGLGGRAQLAMALFEQEQA NST

**Signal sequence:**

amino acids 1-15

**Transmembrane domain:**

amino acids 489-507

**N-glycosylation site.**

amino acids 121-125, 342-346

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 319-323, 464-468

**Casein kinase II phosphorylation site.**

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,  
399-403, 409-413, 473-477, 729-733, 748-752

**Tyrosine kinase phosphorylation site.**

amino acids 736-743

**N-myristoylation site.**

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,  
558-564, 651-657, 657-663, 672-678

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 14-25

**Cell attachment sequence.**

amino acids 247-250

## FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC  
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG  
AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA  
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA  
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAACC  
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG  
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTGTTGAGTCAATTAATATGGACACAAATGACATG  
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACCTG  
GTTCTTCCTTGACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTAA  
AAAAGGATCCATCACAGCCTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT  
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT  
CAATATCCCAGAAAAGTGTCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC  
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA  
AAAGATGTATTTAATACCAAATCTGTTGGGCTTCTATTAAAGAGGCAATGACTTATCACCC  
CAACCAGGTAGTAGAAGGCTGTTGTTGAGATATGGCTGTTACTTTTAATGGACTGACTCCAA  
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTGCGGCATATTTTCAAT  
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG  
TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACATACATATCCAA  
TACAGCTGTATGTTTCTTTTTCTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAG  
TCAGTAGTACATTTTAAATGAGGGTGGTTTTTTCTTTTAAACACATGAACATTGTAAATG  
TGTTGGAAAGAAGTGTTTTAAGAATAATAATTTGCAAATAAACTATTAATAAATATTATAT  
GTGATAAATCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT  
AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT  
TGTGATTAAAGTAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT  
AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAATGTACAATAAGCAACTAAAGAAA  
ATTAAAGTGAAAGTTGAAAAAT

## **FIGURE 120**

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME  
LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK  
AYKYAFDKYRDQYNWFFLARPTTFAI IENLKYFLKKDPSQPFYLGHTIKSGDLEYVGMEGG  
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT  
KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL  
PPNGSDND

### **Signal sequence:**

amino acids 1-33

### **N-glycosylation site.**

amino acids 121-125, 342-346

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 319-323, 464-468

### **Casein kinase II phosphorylation site.**

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,  
399-403, 409-413, 473-477, 729-733, 748-752

### **Tyrosine kinase phosphorylation site.**

amino acids 736-743

### **N-myristoylation site.**

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,  
558-564, 651-657, 657-663, 672-672

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 14-25

### **Cell attachment sequence.**

amino acids 247-250

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 121

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG  
AAAAAGAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAGAAAAAATGAATTCATCTAAATCAT  
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT  
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTTCGCAT  
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTACAGAGCTCTCCTGCT  
ACAATTATGGATCAGGTTCAAGAAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC  
AGCTGCTACTTCTTTTCTACTGACACCATTTCTGCGGCTTAAGTTTAAAGAACTGCTCAGC  
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCTTACAAGA  
AACCTAAAATGAGAGAGTTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA  
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA  
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA  
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG  
AACAAAGGAAAATCTCTTTAAGAACAGAAGGCACAACCTCAAATGTGTAAAGAAGGAAGAGCA  
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC  
ATAAGTATTTGTTACTCTGATACAAATAAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAAAAA  
AAA  
AAAAA

## **FIGURE 122**

MNSSKSSETQCTERGCFSQMFLLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN  
FTELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALS LKNCSAMGAHLVVINSQEEQ  
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS  
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

### **Signal sequence:**

amino acids 1-42

### **N-glycosylation site.**

amino acids 2-6, 62-66, 107-111

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

### **N-myristoylation site.**

amino acids 15-21, 74-80, 155-161

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 27-38

1. The first part of the document is a list of references. The references are listed in a vertical column on the left side of the page. The references are:

1. The first part of the document is a list of references. The references are listed in a vertical column on the left side of the page. The references are:

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCTCAGCAACCCCTCGACATGGCGCTGAGGCGGCCACCGCGAC  
TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCCCTGCTGCTGCTTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC  
TCAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGTGGAAGTGTCTTGCATCATTACGGATTTCGC  
AGACAAGTGACCCAGGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTGACAAACAAA  
TTCAGGGAGACTTGGCGGGTCTGTCAGAAATCTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG  
ACTGACCCCTTTATCGCTGTGAGGTCGTGCTCGAATGACCCGCAAGGAATGTAGATGTGATCGAGTTAA  
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGATGCCCAGAAGCTGTACCAGTAGGCAAGATGGCAACACTGC  
ACTGCCAGGAGAGTGAGGGCCACCCCGGCCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCACCGATT  
CCAGAGCCAATCCAGATTTTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCAGTGCTG  
TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG  
AGATGGAAGTCTATGACCTGAACATTGGCGGAATTTATTTGGGGGGTCTTGGTTGTCTTGTCTGTACTGGCCCTGA  
TCAGTTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTTCATCACTAATAACAGGATGGAGAAAGTTACAAGA  
ACCCAGGGAAACCAGATGGAGTTAACTACATCCGCACGTACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTTG  
TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA  
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTGAGAAGCTTTTCGTTTTGGCCAAAGTTGACCA  
CTACTCTTCTTACTCTAACAAGCCACATGAATAGAAGAAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA  
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTTCTGGCCTGATTTCCCGCATGAGTATTAGG  
GTGATCTTAAAGAGTTTGTCTACGTAAACGCCCGTGTGGGCCCTGTGAAGCCAGCATGTTCCACTGGTGGT  
CAGCAGCCACGACAGCACCATGTGAGATGGCGAGTGGCTGGACAGCAGCCAGCAGCATCCCGCGGGAAACCCA  
GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCACAGACACCACCGAGTTTTCTTCTTAAAGGCTCTGC  
TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTTGGATCAGCATTTTGTAAAAACAACCAAATCAGGAAG  
GTAAATTGGTTGCTGGAAGAGGGATCTTGCTGAGGAACCTGCTTGTCCAACAGGGTGTGAGGATTTAAGGAAA  
ACCTTCGCTTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTGTTTTTATAAAATTT  
TACATCTAAATTTTTTGCTAAGAGTGTATTTGATTTATTTGAAAGAAATTTCTATTTAACTGTAATATATTGT  
CATACAATGTTAAATAACCTATTTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTAAAT  
TGGAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTTCTTTTCT  
CACACAAGTTTTAGCCTTTTTTACAAGGGAACCTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT  
TAAAAATCCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAGAAACCTCTCAGGTTAGCTTTGAACT  
GCCCTTCTCAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGGAAGCCCTCAGATGTACATACAGATG  
CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCCGTCTAGCTCACTGTTGCCCTCGCTGTCTGCCAGGAGCCCT  
GCCATCCTTGGGCCCTGGCAGTGGCTGTGTCCCAGTGAGCTTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC  
TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAAACAGACCTCT  
TTTTGGTTATGGATGGCTCAAAAATAGGGCCCCCAATGCTATTTTTTTTTTTTTAAGTTTGTTTAATTATTGT  
AAGATTGTCTAAGGCCAAAGGCAATGGCGAAATCAAGTCTGTCAAGTACAATAACACTTTTTTAAAGAAAATGGAT  
CCAGTGTCTCTCTTTGGCACAAGAAAGCACCCAGACGCCACAGGCTCTGTGCAATTTCAAACAAAACCATGAT  
GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTGAGGTGGAGCAGCCAGGTGAAGAGCCTGGCGGGAGGAAAG  
TGAAACGCCCTGAATCAAAGCAGTTTTTCTAATTTTGACTTTAAATTTTTTCATCCGCCGGAGACACTGCTCCCATT  
TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT  
GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCTGGA  
GAATGGCTCTCACTACTCACTTGTCTTTTCACTTCCAGTGTCTTGGGTTTTTTTATACATTTGACAGCTTTTTTTT  
AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTATGTGAACACTTTGCCCGAGGCCCTGGCAGAGGCA  
GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTCAGCATGGCATCCCTGGATGCTTAGCATGCAAGTTC  
CCTCCATCATTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTACGCTCCAGCCTCCT  
CTTGGTTGTGATAGTATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCTTAAACCTTCTACACTAGTGCCA  
TGGGAAACCAGGTCTGAAAAGTAGAGAGAAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA  
CGGAAATAGGAATACCTGTGTATTTTAAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT  
GCCTTTGGATGGATGTTGTGTTTACACAGATGCTACAGACTTGTAACACACCGTAATTTGGCATTGTGTTAAC  
CTCATTTATAAAAGCTTCAAAAAACCCA

## FIGURE 124

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624

><subunit 1 of 1, 310 aa, 1 stop

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS  
PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTS�KIWNVTRRDSALYRCEVVARNDRK  
EIDEIVIELTVQVKPVPVPCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRA  
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDLNIIGGIIGG  
VLVVLAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-30

#### **Transmembrane domain:**

amino acids 243-263

#### **N-glycosylation sites.**

amino acids 104-107, 192-195

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 107-110

#### **Casein kinase II phosphorylation site.**

amino acids 106-109, 296-299

#### **Tyrosine kinase phosphorylation site.**

amino acids 69-77

#### **N-myristoylation sites.**

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267